SPEECH & HEARING CLINIC
CLARION STATE COLLEGE

JOURNAL

OF SPEECH AND HEARING

RESEARCH

VOLUME 1, NUMBER 3 SEPTEMBER, 1958

A Journal of the
AMERICAN SPEECH AND HEARING ASSOCIATION

MANUSCRIPTS and related correspondence should be addressed to Dorothy Sherman, Editor, *Journal of Speech and Hearing Research*, Department of Speech Pathology and Audiology, East Hall, State University of Iowa, Iowa City, Iowa.

SUBSCRIPTIONS and orders for back numbers should be addressed to Kenneth O. Johnson, Business Manager, *Journal of Speech and Hearing Research*, American Speech and Hearing Association, 1001 Connecticut Avenue N. W., Washington 6, D.C.

NOTICE TO AUTHORS

Before submitting manuscripts for publication authors should consult Information for Contributors to the Journal of Speech and Hearing Research, *JSHR*, 1, 1958, 94-96. Prospective authors are invited to write the Editor for copies of this Note, and for its supplement, Examples of Manuscript Form. It is assumed that authors keep carbon copies of their manuscripts since the *JSHR* cannot assume responsibility for loss. No manuscript can be considered for publication in the *JSHR* which has appeared elsewhere.

The Association reserves the right to keep all copies of any manuscript submitted to it.

NOTICE TO MEMBERS

Member authors of ASHA are requested to provide autographed copies of their bound publications for a library in the National Office. Books should be sent to Kenneth O. Johnson, Executive Secretary, American Speech and Hearing Association, 1001 Connecticut Avenue, N.W., Washington 6, D. C.

The *Journal of Speech and Hearing Research* is published in March, June, September and December of each year at the Interstate Printers & Publishers, Danville, Illinois.

Anyone has permission to reproduce a portion of any article in this issue of the *Journal* if permission is also obtained from the author. Quotations must be accurate, and credit must be given to the author or authors. Credit must be given to the *Journal of Speech and Hearing Research* by bibliographic reference or other suitable form.

Subscription: $5.00 per year in U.S.A.; $5.50 per year outside U.S.A.; $1.35 per single copy. Entered as second-class matter at the Post Office in Danville, Illinois, under Act of Congress of March 3, 1879. Printed by The Interstate Printers and Publishers, Inc., 19-27 North Jackson Street, Danville, Illinois. Office of Publication, Jackson at Van Buren, Danville, Illinois.

SPEECH & HEARING CLINIC
CLARION STATE COLLEGE

JOURNAL OF SPEECH AND HEARING RESEARCH

Volume 1 **September 1958** **Number 3**

Posterior Pharyngeal Wall Movement in Normals
ROBERT F. HAGERTY, MILTON J. HILL, HAROLD S. PETTIT AND JOHN J. KANE 203

GSR Auditory Threshold Mechanisms: Effect of Tonal Intensity on Amplitude and Latency under Two Tone-Shock Intervals
A. E. ARONSON, J. E. HIND AND J. V. IRWIN 211

GSR Auditory Threshold Mechanisms: Instrumentation, Spontaneous Response and Threshold Definition
J. E. HIND, A. E. ARONSON AND J. V. IRWIN 220

Myoelastic-Aerodynamic Theory of Voice Production
JANWILLEM VAN DEN BERG 227

Infant Speech: Consistency with Age
HARRIS WINITZ AND ORVIS C. IRWIN 245

Syllabic and Phonetic Structure of Infants' Early Words
HARRIS WINITZ AND ORVIS C. IRWIN 250

Application of Descriptive Linguistics to Child Language
ROBERT W. ALBRIGHT AND JOY BUCK ALBRIGHT 257

Auditory Skills of Blinded Individuals Training with Pilot Dogs
JOHN J. O'NEILL, HERBERT J. OYER AND DONALD J. BAKER 262

Mental Imagery
CHARLES F. DIEHL AND NOEL C. ENGLAND 268

Traditional and Psychogalvanic Skin Response Audiometry
KENNETH W. BURK 275

Language Skills in Children with Cleft Palates
D. C. SPRIESTERSBACH, FREDERIC L. DARLEY AND HUGHLETT L. MORRIS 279

Skin Resistance Changes Accompanying the Sidetone Test for Auditory Malingering
CLAIR N. HANLEY, WILLIAM R. TIFFANY AND JACQUELINE M. BRUNGARD 286

Behavioral Rigidity in Adult Stutterers
ERVIN KAPOS AND LLOYD S. STANDLEE 294

THE AMERICAN SPEECH AND HEARING ASSOCIATION

OFFICERS

President	Jon Eisenson, Ph.D., Queens College
Executive Vice-President	Stanley Ainsworth, Ph.D., Univ. of Georgia
Vice-President	Leo G. Doerfler, Ph.D., Univ. of Pittsburgh
Editor of the Association	Robert West, Ph.D., Brooklyn College

OFFICERS-ELECT

President-Elect	George A. Kopp, Ph.D., Wayne State Univ.
Executive Vice-President-Elect	Jack Matthews, Ph.D., University of Pittsburgh
Vice-President-Elect	Miriam D. Pauls, Ph.D., Johns Hopkins Med. Instit.
Editor of the Association-Elect	Wendell Johnson, Ph.D., Univ. of Iowa

COUNCIL

The Officers and the following Councilors:

Raymond Carhart, Ph.D.
Ernest H. Henrikson, Ph.D., (1956-58)
Ira J. Hirsh, Ph.D., (1958-60)
Ruth B. Irwin, Ph.D., (1956-59)
Herold S. Lillywhite, Ph.D., (1958-59)
Hayes A. Newby, Ph.D., (1957-59)
Wilbert L. Provonost, Ph.D., (1958-60)
Sylvia O. Richardson, M.D., (1957-59)
Charlotte G. Wells, Ph.D., (1956-58)

EXECUTIVE SECRETARY

Kenneth O. Johnson, Ph.D.

APPLICATIONS FOR MEMBERSHIP SHOULD BE ADDRESSED TO THE EXECUTIVE SECRETARY

THE JOURNAL OF SPEECH AND HEARING RESEARCH

EDITOR

Dorothy Sherman, Ph.D.

ASSISTANT TO THE EDITOR

Jean B. Kern, Ph.D.

ASSISTANT EDITOR

Frederic L. Darley, Ph.D.

STATISTICAL CONSULTANT

Leonard S. Feldt, Ph.D.

ASSOCIATE EDITORS

Oliver Bloodstein, Ph.D.
Arthur S. House, Ph.D.
James F. Jerger, Ph.D.
D. E. Morley, Ph.D.
Hildred Schuell, Ph.D.
William R. Tiffany, Ph.D.
John C. Webster, Ph.D.
Joseph M. Wepman, Ph.D.

DEPARTMENT EDITORS

Ernest H. Henrikson, Ph.D. *Book Reviews*
Martin F. Palmer, Sc.D. *Records*

EDITOR OF THE ASSOCIATION

Robert West, Ph.D.

BUSINESS MANAGER

Kenneth O. Johnson, Ph.D.

Posterior Pharyngeal Wall Movement In Normals

Robert F. Hagerty

Milton J. Hill

Harold S. Pettit

John J. Kane

For almost 100 years the complex inter-activities of the structures controlling the velopharyngeal port have stimulated conjecture, interest, and controversy among members of the several disciplines which concern themselves with this function. Foremost among these have been anatomists, surgeons, speech therapists, dental specialists, and radiologists. During the past quarter century particular attention has been paid to velopharyngeal muscle function by those engaged in the habilitation of cleft palate persons, because good speech is considered to be (1) the primary goal in treatment of these individuals and (2) more directly related to adequate velopharyngeal closure than to any of the other anatomical differences found in individuals with cleft palate. The present paper is designed to shed light on only one of the multiple aspects of velopharyngeal closure—the forward excursion of the posterior pharyngeal wall during normal production of certain speech sounds.

Robert F. Hagerty (M.D., Johns Hopkins University, 1943) is Assistant Professor of Surgery (Plastic Surgery), Medical College of South Carolina. Milton J. Hill (M.S., Pennsylvania State University, 1954) is Instructor in Speech Therapy and Audiology, Medical College of South Carolina. Harold S. Pettit (M.D., Western Reserve University, 1939) is Professor of Radiology, Medical College of South Carolina. John J. Kane (M.D., Medical College of South Carolina, 1938) is Instructor in Radiology, Medical College of South Carolina. This investigation was carried out at the Cleft Lip and Cleft Palate Center of the Medical College Hospital, and was supported by a research grant (D-197) from the National Institute for Dental Research, of the National Institutes of Health, Public Health Service.

During the 1860's, Passavant (*6*) described muscular activity which he had observed in the posterior pharyngeal walls of some of his patients, as follows:

> . . . the lower part of the pharynx is raised while the upper part of the posterior pharyngeal wall remains immovably in place. In addition, a transverse swelling arises on the posterior wall of the pharynx at the level of the bases of the uvular halves. From this transverse swell-

> ing, folds run upward and outward; not infrequently fold formation also occurs at the lower border of the swelling, running downward and outward. While the posterior pharyngeal wall is thus swollen, an approach of the two lateral pharyngeal walls toward each other occurs at the same level. This transverse swelling, which becomes visible on the posterior pharyngeal wall somewhat above the uvular halves, is of the greatest importance for speech.

In the cleft palate subject, movement in the posterior pharyngeal wall, if it takes place, is readily observable. In the subject having a normal palate such movement defies observation because the velum and uvula hang down shielding it. Because of this, the collection of objective data concerning activity in this area during the production of speech sounds by persons with normal palates has only recently become possible through advance in radiological technique.

Although several investigators (*4*, *5*, 7, *8*, *9*) had previously used x-ray techniques in studying the action of the velum and posterior pharyngeal wall, the investigation by Wolfe (*11*) in 1942 is especially pertinent to the present context. Using x-rays of 15 randomly selected normal subjects he studied 'the amount of movement of the posterior wall of the pharynx, the relative position of the bulge at the back wall, the amount and nature of the movement of the velum and the position of the velum, in relation to the back wall of the pharynx during phonation' of sustained vowels. He established a nasal reference line (drawn between the anterior end of the nasal spine and a point at the junction of the hard and soft palate) as his basic reference for the purpose of making tracings of his films. This served as a base line for the construction of all perpendiculars and parallels needed for making the desired measurements.

In 1949 Hixon (*2*), in an x-ray study of oral and pharyngeal structures, included among his data measurements of forward excursion of the posterior pharyngeal wall, but he stated that a 'mottling' or 'graininess' of the film obliterated some of the detail in that area. Williams (*10*), in an investigation of 30 normal subjects in 1952, made measurements on the bases of perpendiculars erected from the Frankfort plane. In 1954 Buck (*1*), using Wolfe's technique for obtaining the measurements, compared the velopharyngeal movements of 20 postoperative cleft palate subjects with those of 20 normals. His findings with respect to posterior wall movement for the normals agreed in general with those of Wolfe and Hixon.

The present study was undertaken in order to establish an objective set of normative data with which to compare measurements of postoperative cleft palate subjects. Three crucial differences exist between this and previous studies of similar nature. First, the cephalometric laminagraphs were made with the subjects in an upright position rather than supine as has been the case in previous laminagraphy. Second, data were collected on a large number of subjects whereas previous numbers on which data have been reported range from three to 30. Third, the continuant consonant [s] was used as one of the subject conditions. Previous studies have not included data on consonant sounds.

Procedures and Equipment

The laminagraphic x-ray technique was used in order to assure that the

tracings which were being made from the skull films would depict more accurately the midline contours of the structures being measured. Conventional x-rays do not afford clear soft tissue outlines since the image obtained is a summation of shadows of tissue masses. Superimposition of tissue structures is a handicap in obtaining a radiograph which will show the most diagnostic information. Laminagraphy offers a means of selecting a desired 'layer' or plane of tissue structures by blurring out tissues on either side of the selected plane. With the intention of facilitating tracing of the films, a radiopaque substance, Dionosil, was used with each subject.

The data herein reported were derived from measurements of lateral upright cephalometric laminagrams of subjects with normal palates and no apparent skull malformations. Eighty persons, 50 whites and 30 Negroes associated with the Medical College and Hospital of South Carolina, volunteered to cooperate with the investigators. It was assumed that there might be differences between the two groups, so the data on Negroes and whites were treated separately. The age range of the white subjects was 19 to 62 years, with a mean age of 27.3 years. The age range of the Negro subjects was 15 to 75 years, with a mean age of 44.8 years. The subjects were carefully prepared for maintaining the conditions desired during the time the x-rays were being taken, and trained personnel were present to listen and judge whether the sounds being produced were those which were desired. The laminagraphs were taken under three conditions: with the soft structures at rest; during production of the vowel [ɑ][1]; and during production of the continuant [s].

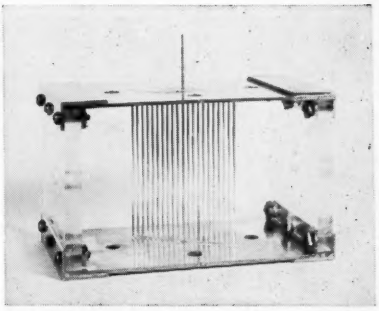

FIGURE 1. Photograph of laminometer.

The procedure for obtaining the upright cephalometric laminagraphs was as follows: the subject was first placed in a supine position and 1 cc of water-based Dionosil was instilled into each nostril. After about 30 seconds the subject was allowed to sit erect.[2] He was then seated on a stool of adjustable height and his head immobilized in true lateral position by the cephalostat. The true lateral position was corroborated in each instance by conventional x-rays (taken previous to the laminagraphs) showing superimposition of the brass bands about the ear rods.

The thickness and position of the 'cut' were corroborated by x-rays of

[1]The vowel sound which, in General American speech, is heard in the words *top, hop, pot* and *mop.*

[2]An oil-based Dionosil was later substituted which required five minutes in the supine position for proper distribution. This was done because subjects complained of a stinging, burning sensation with the water-based material. The results were equally satisfactory.

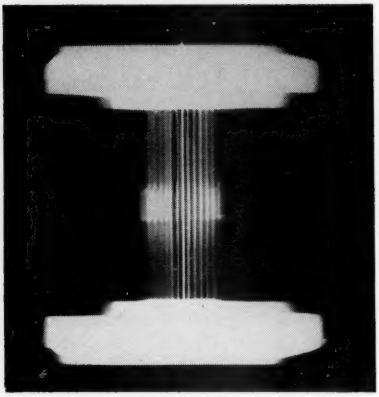

FIGURE 2. Laminagraph of laminometer. The sharply outlined central wire indicates proper midline focus.

a laminometer in place in the cephalostat. The laminometer consists of multiple parallel wires set 2 mm aside and 1 mm in front of each other in an oblique line (Figure 1). The longer central wire extends through one of the two metal plates which are firmly secured to the plastic endpieces. The centers of these plastic endpieces are drilled to receive the ear plugs. After the proper adjustments are made, an x-ray taken with this device in place shows the 'cut' to be through the long central wire which is in clear outline, and the thickness of the plane in focus extends to no more than the first wire on either side of the midline (Figure 2), giving a thickness of 'cut' of *ca* 4 mm or less.

The x-ray equipment used is a modified General Electric Ordograph (laminagraph) unit to which is attached a cephalostat. The laminagraph consists of a moving bucky grid interlocked with the x-ray tube by means of an arm so that the desired plane running through the center of the skull is at the fulcrum point or axis of the arm. The film moves on a track in the cassette holder over a range of 10.64 cm. The cephalostat is attached to this cassette holder so that a true lateral projection of the skull is assured.

The tube stand, which is floor mounted and hydraulically driven, moves through a range of 53.34 cm, the exposure being made during the 38.1 central cm of travel. The distance from the tube target to the film is 95.25 cm in the mid-position. The distance from the midline of the skull to the film is 15.88 cm in the lateral position. The magnification of the image on the film is 12.1%.

Many combinations of x-ray factors were used in an effort to achieve optimal contrast of air, soft tissue, and bone in the radiograph. It was finally decided that the best results were obtained with 65 kv, 10 ma and 3 sec for subjects whose width of skull was 14.5 cm or less. The kilo-

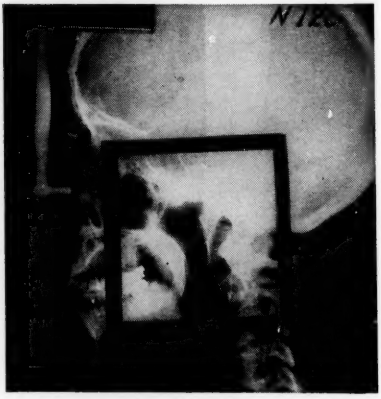

FIGURE 3. Conventional radiographic skull film of subject in rest position. Square represents areas included in Figures 4, 5 and 6.

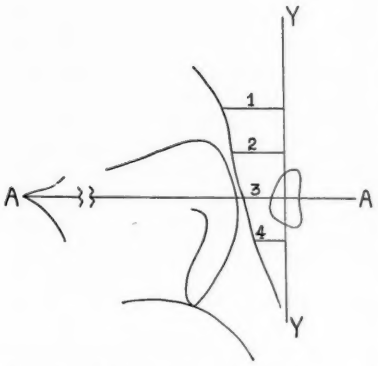

FIGURE 4. Line drawing of laminagram of subject phonating [ɑ]. Drawing shows main reference line, A-A, drawn from anterior nasal spine to midpoint of the tubercle of the atlas; the perpendicular Y-Y, erected at the midpoint of the tubercle; and the levels 1, 2, 3 and 4, at which the distance from Y-Y to the posterior pharyngeal wall was measured.

voltage was increased by five for subjects whose measurements were greater.

All x-ray films were developed in an automatic unit to assure consistent results. Tracings were made of the portions of the films within the inset shown in Figure 3, and the main reference line, A-A (Figure 4), was drawn from the anterior nasal spine through the midpoint of the tubercle of the atlas as suggested by Krogman (*3*). A perpendicular to line A-A was erected at the midpoint of the tubercle and the distance from this perpendicular, Y-Y, to the surface of the posterior pharyngeal wall was measured at four arbitrarily selected levels: (1) along a line 2 cm above A-A and parallel to it; (2) along a line 1 cm above A-A and parallel to it; (3) along A-A from the midpoint of the tubercle, and (4) on a line 1 cm below A-A and parallel to it. Subsequent examination of the tracings disclosed that the actual point of greatest excursion occurred much more frequently between two levels (where it missed being measured) than at a given level. In an attempt to evaluate further the amount of posterior pharyngeal wall movement, a Precision Rolling Disc Planimeter was used to measure the area encompassed by lines 1 and 4, Y-Y, and the posterior pharyngeal wall. Thus, five measurements were taken from the tracing of each laminagraphic skull film, so that 15 measurements were recorded for each of the 80 subjects studied.

Results

Among the white subjects the mean forward movement of the posterior pharyngeal wall from the rest position to [ɑ] at level 1 was .19 mm; at level 2 it was .67 mm; level 3, .70 mm; level 4, .64 mm. From rest to [s] the mean amount of forward excursion at level 1 was 1.57 mm; level 2, .66 mm; level 3, .32 mm; level 4, .04 mm.

Planimetry disclosed greater area for [ɑ] than for rest within the defined confines in 31 of the 50 white subjects. No change was seen in seven, while 12 exhibited less area. The mean for the 50 subjects was .176 sq cm greater for [ɑ] than rest.

Similarly, 31 subjects presented greater areas for [s] than for rest (not all the same subjects, however) while six subjects showed no change and in 13 the area decreased. The average for the 50 here was .188 sq cm more area during production of [s] than in the rest position.

The mean forward movement of the posterior wall of the Negro subjects from the rest position to [ɑ] at level 1 was .9 mm; at level 2 it was 1.733 mm; level 3, 1.233 mm; level 4, .95 mm. From rest to [s] the amount of forward excursion at level 1 was 2.05 mm; level 2, 1.20 mm; level 3, .833 mm; and level 4, 1.333 mm. Of the 30 Negro subjects, 23 presented more area for [ɑ] than for rest within the defined limits measured by planimetry, while seven decreased in area. The average increase in area from rest to [ɑ] in the 30 subjects was .270 sq cm. From rest to [s] 18 increased and 12 decreased, with a mean of .325 sq cm greater area for [s] than rest.

The average forward movement of the pharyngeal wall at each level, and the average increase in area by planimetry was greater in the Negro group in each instance. The differences between the averages for the groups were as follows: rest to [ɑ], level 1, .71 mm; level 2, 1.06 mm; level 3, .53 mm; level 4, .31 mm; planimetry, .094 sq cm. Rest to [s], level 1, .48 mm; level 2, .54 mm; level 3, .51 mm; level 4, 1.29 mm; planimetry, .14 sq cm. Each of these figures is the amount by which the mean for the Negro group exceeded the mean for the white group. It would appear that there was ample justification for considering the two groups separately.

Discussion

Although it is quite possible that statistically significant differences would emerge from further treatment of the data, it is doubtful whether the actual extent of excursion is sufficient to make important differences in speech except when the individual possesses a Passavant's bar large and vigorous enough to participate actively in closure.

Comparison of the measurements of excursion during the production of [ɑ] and [s] reveals less movement in

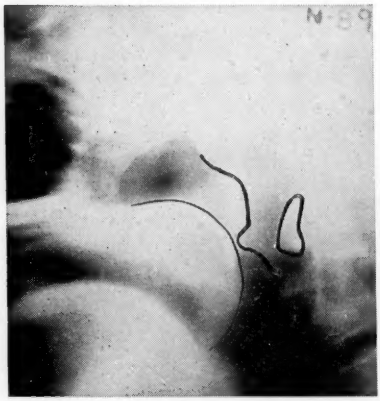

FIGURE 5. Laminagraph showing Passavant's bar in subject phonating [ɑ]. The soft tissues have been outlined in certain areas to provide for greater clarity in the photographic reproduction.

FIGURE 6. Laminagraph showing Passavant's bar in subject saying [s]. The soft tissues have been outlined in certain areas to provide for better clarity in the photographic reproduction.

general for [s] than for [ɑ]. In many of the films of subjects saying [s], while strong velopharyngeal closure is taking place, the velum is seen to impose its contours upon the pharyngeal wall, so that its convex posterior surface creates a corresponding indentation in the pharyngeal wall where the two are in contact. Velopharyngeal contact is seen during production of [ɑ] in 16 of the Negro and 11 of the white subjects, but the extent of this contact is not as great as in [s], and it does not appear that as much force is exerted for closure during [ɑ] as for [s] because the angle the velum shows is less acute.

For their purposes the authors defined Passavant's bar as an obvious forward bunching of the posterior pharyngeal wall within a restricted range. Nine of the 80 subjects studied showed Passavant's bar (see Figures 5 and 6). In three of these this ridge was seen only in the laminagraph of the subject saying [s], and the area of velar contact with the pharyngeal wall in these three subjects was above the ridge. In three others of the group of nine, Passavant's bar was seen during production of the vowel [ɑ], but closure was so vigorous in [s] that the bar again appeared unnecessary if it indeed continued to exist. The laminagraphs of the remaining three subjects in whom Passavant's bar was seen show that it was an active factor in constriction of the port for production of [ɑ], and closure for [s] would have been unsuccessful had no bar arisen.

Anatomists, surgeons, radiologists, dental specialists and speech therapists have long engaged in conjecture and controversy with respect to the participation of the posterior wall of the pharynx in normal speech. The present study was one more attempt to clarify this aspect of the complex activity in the velopharyngeal area. It differs from previous studies in that the cephalometric laminagraphs were made with the subjects in an upright position, a much larger number of subjects was studied, and a consonant, [s], was included as one of the subject conditions.

Summary

Upright cephalometric laminagraphs were made of 80 normal subjects (50 white and 30 Negro) at rest, phonating [ɑ], and producing the continuant consonant [s]. Tracings were made of the laminagraphs and a reference line was established from the anterior nasal spine to the midpoint of the tubercle of the atlas. A perpendicular to this line was erected at the midpoint of the tubercle, and measurements were made from this perpendicular to the posterior wall of the pharynx at four arbitrarily selected levels: 1, 2 cm above the midpoint of the tubercle; 2, 1 cm above; 3, at the level of the midpoint; and 4, 1 cm below. White and Negro subjects were treated as separate groups because of possible anatomical differences. Planimetric measurements were also recorded to determine increases in area as a result of forward movement.

Passavant's bar was evident in the x-rays of nine of the 80 subjects, but in six of the nine actual velar contact with the posterior pharyngeal wall was made above the ridge and so strongly that it appeared Passavant's bar was not necessary for functional closure. In the remaining three subjects the only velopharyngeal contact was on the apex of the ridge.

In general, except when Passavant's phenomenon occurs as a function necessary to closure it is doubtful whether the actual extent of forward excursion of the posterior pharyngeal wall is clinically significant in the production of speech sounds. It remains to be seen whether cleft palate individuals present a similar pharyngeal wall profile.

References

1. Buck, M., Post-operative velo-pharyngeal movements in cleft palate cases. *JSHD,* 19, 1954, 288-294.
2. Hixon, E., An X-ray Study Comparing Oral and Pharyngeal Structures of Individuals with Nasal Voices and Individuals with Superior Voices. M. S. Thesis, State University of Iowa, 1949.
3. Krogman, W. M., Personal communication to the authors.
4. Norris, M. A., X-ray Studies of Vowel Production as it is Related to Voice. M.A. Thesis, State University of Iowa, 1934.
5. Parmenter, C. E. and Trevino, S. N., Vowel positions as shown by x-ray. *Quart. J. Speech,* 18, 1932, 351-369.
6. Passavant, G., Über die Verschliessung des Schlundes beim Sprechen. *Virchows Arch.,* Bd. XLVI (Vierte Folge Bd. VI), Hft. 1, 1-32.
7. Russell, G. O., First preliminary x-ray consonant study. *J. acoust. Soc. Amer.,* 5, 1934, 247-251.
8. Scheier, M., Die Verwerthung der Röntgen-Strahlen für die Physiologie der Sprache und Stimme. *Archiv. für Laryngologie und Rhinol. Berlin,* 7, 1897, 116-127.
9. Scheier, M., Die Anwendung der X-Strahlen für die Physiologie der Stimme und Sprache. *Verhandl. d. Cong. f. innere Med. Wiesbaden,* 15, 1897, 531-533.
10. Williams, R. L., A serial radiographic study of velopharyngeal closure and tongue positions in certain vowel sounds. *Northwestern Univ. Bull.,* 52, 17, 9-12.
11. Wolfe, W. G., X-ray Study of Certain Structures and Movements Involved in Naso-Pharyngeal Closure. M.A. Thesis, State University of Iowa, 1942.

GSR Auditory Threshold Mechanisms: Effect Of Tonal Intensity On Amplitude And Latency Under Two Tone-Shock Intervals

A. E. Aronson

J. E. Hind

J. V. Irwin

Several difficulties are inherent in the GSR or, to use a recently suggested term (2), the EDR (electrodermal response) audiometric technique. One is the frequent occurence of spontaneous responses caused by deep breaths, limb movements and alterations in thought processes. The simultaneous occurrence of such activity with a tone will result in an EDR that can be easily mistaken for a valid response.

Clinicians rely primarily upon a latency criterion for differentiating between valid and spontaneous EDR's. An arbitrarily established latency range of approximately 1.0 to 3.5 seconds provides a frame of reference for making this distinction. EDR's falling within these limits are accepted as valid responses while those falling outside these limits are rejected as spontaneous. While Stewart (*10*) has shown for a group of subjects that EDR latencies to tones could be distinguished from responses to controls with a high degree of certainty, he also discovered that for individual subjects this differentiation could not be made three out of 15 times on the average.

EDR amplitude has played almost no role in contributing to the differentiation between valid and spontaneous EDR's. Although a minimum height criterion is often imposed below which an apparent change in resistance or potential is not accepted as valid, these amplitude criteria are nevertheless arbitrarily chosen and the response heights influenced by the physical characteristics of the recording instrument. The inadequacy of existing latency and amplitude criteria stems from a lack of precise in-

A. E. Aronson (Ph.D., Wisconsin, 1957) is Assistant Professor of Speech, University of Wisconsin; J. E. Hind (Ph.D., Chicago, 1952) is Assistant Professor of Physiology in the Laboratory of Neurophysiology, University of Wisconsin; and J. V. Irwin (Ph.D., Wisconsin, 1947) is Professor of Speech and Director of the Speech and Hearing Clinic, University of Wisconsin. This study was aided by grants from the Alfred Laukhuff Trust for research in the field of cerebral palsy, from the National Institute of Neurological Diseases and Blindness (B-896) and from the Research Committee of the University of Wisconsin out of funds supplied by the Wisconsin Alumni Research Foundation.

formation on how EDR latency and amplitude change quantitatively as tonal intensity changes from below-threshold to above-threshold levels. That such changes do occur has been demonstrated by Davis (*1*).

Difficulty in conditioning presents a second important problem in EDR audiometry. Neurologically involved subjects, the hearing handicapped, and even normal subjects have shown themselves to be difficult to condition (*3, 5, 9*). The problem of obtaining high strength of conditioning in EDR audiometry has been investigated by Meritser and Doerfler (7) who demonstrated the superiority of 40% reinforcement over 100% reinforcement. Schlosberg and White (*8*) have shown that 0.5 second is the optimal CS-US interval for obtaining the largest EDR's to light stimuli, suggesting the application of a similar CS-US interval to EDR audiometry.

With the above problems in mind, the following factors were investigated: (1) the effect of tones presented at subthreshold, threshold and above-threshold intensities on (a) EDR latency, (b) variability of EDR latencies and (c) EDR amplitude; (2) the effects of conditioning the EDR to a short tone of 0.6 second followed instantaneously by shock instead of the more widely used method of a long tone of 3.0 seconds followed instantaneously by shock, on (a) strength of conditioning as measured by EDR amplitude and number of successfully conditioned subjects, (b) variability of EDR latencies to tones presented at sub-threshold, threshold and above-threshold intensities and (c) rate of growth of EDR amplitudes to tones presented at sub-threshold, threshold and above-threshold intensities; (3) the comparison between the resistance and the potential methods of recording EDR's.

Method

Subjects and Conditions. Forty men and women from University of Wisconsin classes in the Department of Speech were divided into two groups of 20 each. Both groups were submitted to the following procedures: (1) a voluntary threshold test, (2) EDR conditioning and (3) an EDR threshold test.[1]

Tones of 1000 cps and 0.6 second duration were used to find the voluntary thresholds for all subjects. However, for EDR conditioning and the EDR threshold test, although 20 subjects, designated the 'short tone' group, received tones of 0.6 second duration, the remaining 20 subjects, designated the 'long tone' group, received tones of 3.0 seconds duration. For the EDR conditioning and EDR threshold phases, programmed schedules of unreinforced and reinforced tones were used. Random 40% reinforcement was employed in the conditioning phase and 37.5% reinforcement in the EDR threshold test. The shocks were presented immediately following the ends of the tones. Thus, the 0.6 second and 3.0 seconds values represent the period between the onset of the tone and onset of the shock as well as the duration of the tone. The programs of stimulus pres-

[1]Also administered were five preconditioning trials given after the voluntary threshold test to determine the degree to which each subject gave CR's before being conditioned and 10 extinction trials following the EDR threshold test. The results of these procedures will not be discussed.

entations included intervals of 15, 22 and 30 seconds distributed between trials in random fashion.

In order to minimize the possible effects of the order of tones and reinforcements on the test results, two out of many possible programs of tones and reinforcements were used. Ten subjects, or half of each major group, received conditioning and EDR threshold tests under one of these two arbitrary patterns. The remaining half of each major group followed the second pattern. The assignments of subjects to one or the other of the two patterns were systematically varied according to the order in which the subjects were tested. Because resistance changes were recorded from one hand and potential changes from the other, five subjects from each subgroup of 10 had the resistance electrodes attached to the left hand and the potential electrodes to the right hand, while for the remaining five subjects these attachments were reversed. This modification was introduced in order to counteract the possible effect of laterality on the test results.

Technique and Instrumentation. The EDR recording equipment is described in detail in another article (*4*). Briefly, potential and resistance changes were presented on the vertical axes of a dual beam cathode ray oscilloscope. Because the horizontal sweeps were made to begin simultaneously with the tones, responses of constant latency appeared at the same position on the scope face on successive trials. A Grass kymograph camera was used to register each sweep on a single frame of 35 mm photographic paper. Measurements of response amplitude and latency were made from the photographic records. A second dual beam scope was used for visual monitoring of the responses since the camera was connected through a light-proof hood to the main scope. The horizontal sweeps of the two beams were adjusted to be approximately equal in length, but individual time calibrations were recorded on each trace each day so that the appropriate scale could be used for reaching the potential and conductance latencies. The total duration of each sweep was approximately 7.4 seconds.

Only 1000 cps tones were employed in this study. The rise-time, or time required for the tone intensity to attain its maximum value, and the fall-time, or time required for the intensity to drop to zero, were both approximately 50 msec. The two tone durations used in this study, 0.6 second and 3.0 seconds, were measured from the beginning of their rise to the end of their fall. The tones were presented to the right ear only by way of a calibrated Permoflux type PDR-10S earphone mounted in a semi-plastic earphone socket. The opposite ear was covered by a dead phone with a similar socket.

The shocks were applied to the outer surface of the right leg just below the head of the fibula. The shock electrodes consisted of a pair of one-half inch in diameter steel buttons mounted in a plastic sheet so that their centers were .75 inch apart. A very thin layer of electrocardiograph paste was applied to the buttons before the assembly was attached to the leg by means of an elastic band. Each shock consisted of a 0.4 second train of pulses that began at the instant the tone intensity dropped to zero. The shock current was obtained from

a shielded transformer driven by a laboratory stimulator. The frequency of the individual current pulses in the train was 60 per second; the time constant of each pulse was approximately 1.5 msec. Three different shock levels were used in the experiment. At the appropriate stages in the experiment the stimulator was adjusted for the same three open circuit output voltages for all subjects. Calibration of the actual output current on two subjects yielded values of peak current of 7, 10 and 14 ma, corresponding to the low, medium and high intensity settings.

For each subject simultaneous recordings were made from two electrode pairs on the right and left hands, one pair for the potential changes and the other for the conductance changes. The electrodes used for recording these changes were zinc discs approximately five-eighths inch in diameter with slightly concave surfaces. The hollow sides were filled with a paste composed of saturated zinc sulfate and Bentonite clay. Two electrodes were applied to each hand with strips of masking tape. One electrode was applied to the palmar surface over the second metacarpal-phalangeal joint and the other to the dorsal surface at about the midpoint of the second metacarpal bone.

Procedure. The first test administered to each subject was the voluntary threshold test. The subject was instructed to press a button as soon as he heard the tone. Monaural 1000 cps tones of 0.6 second duration were presented in descending order. Threshold was defined as the lowest intensity level at which the subject responded to two or more out of four trials. Only subjects with thresholds within the range from 0 to −20 db *re* audiometric zero were accepted.

Before entering the EDR conditioning phase, the examiners selected the appropriate tone duration schedule which consisted of either 0.6 second tones or 3.0 second tones. For conditioning, a 1000 cps tone at +10 db *re* voluntary threshold was used. As stated earlier, 40% random reinforcement was employed, and adjacent trials were separated by intervals of 15, 22 and 30 seconds distributed randomly. For the first three reinforcements the shocks were presented at

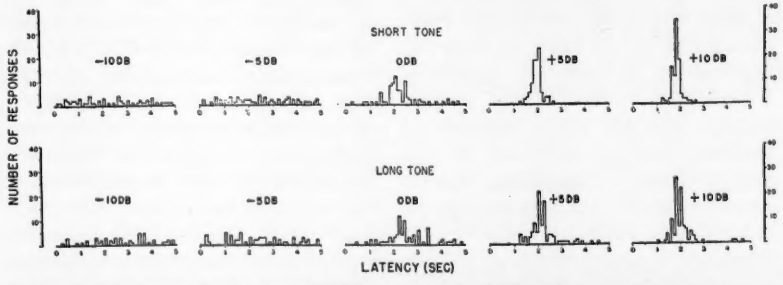

FIGURE 1. Frequency distribution of EDR latencies for long and short tone groups.

TABLE 1. Analysis of variance of EDR magnitudes.

Source	*df*	*ms*	*F**	$F_{.01}$
Between subjects	(39)			
Duration (D)	1	6.66	10.95	7.50
Schedule (S)	1	0.84	1.38	
Hand (H)	1	0.20	0.33	
D x S	1	0.19	0.31	
D x H	1	0.37	0.61	
H x S	1	0.11	0.19	
D x S x H	1	0.77	1.27	
Subjects in Procedure (Error Between)	32	0.61		
Within subjects	(160)			
Intensity (I)	4	11.49	108.43	3.47
D x I	4	1.29	12.20	3.47
I x S	4	0.29	2.69	
I x H	4	0.04	0.38	
I x D x S	4	0.05	0.47	
I x D x H	4	0.09	0.83	
I x S x H	4	0.04	0.38	
I x D x S x H	4	0.13	1.23	
Subjects within x Intensity (Error Within)	128	0.11		
Total	199			

*All between effects tested by mean square for error between; all within effects tested by mean square for error within.

the low intensity setting of the stimulator; the remaining shocks during conditioning were presented at the medium setting. The subject was considered conditioned after responding to seven unreinforced tones that were not necessarily consecutive. If, after a total of 20 trials, this criterion could not be met, the subject was dropped from the sample.

When the subject met the conditioning criterion, his EDR threshold was then determined. This procedure involved presenting a total of 40 unreinforced and reinforced tones at 1000 cps. These trials can be broken down into eight tones at each of the following five intensity levels *re* voluntary threshold: +10 db, +5 db, 0 db, −5 db and −10 db. Since 37.5% reinforcement was used, three out of the eight tones at each intensity level were reinforced, leaving five scorable responses at each of the five intensity levels. The shock intensity was held at the medium setting until half way through the test; at the 21st trial, the shock was elevated to the high setting.

Analysis[2] and Results

First, mean EDR latencies and amplitudes were computed for each

[2]The authors wish to acknowledge the assistance of Dr. David A. Grant and Dr. Richard F. Thompson with the statistical aspects of this study.

subject's responses at each of the five intensity levels. However, because of the fact that on numerous occasions subjects gave no responses to tones, the problem arose of how to treat these 'no responses.' It was decided that mean EDR heights would be derived in two different ways: (1) the sum of the EDR heights in micromhos was divided by the number of *actual responses* and the resulting value called 'mean amplitude'; (2) the sum of the EDR heights in micromhos was divided by the number of *trials* and the resulting value called 'mean magnitude.' Thus, mean amplitude corresponds to the actual mean height seen on the recorder, whereas mean magnitude is a mathematical definition that combines in one expression the factors of amplitude and the percentage of trials on which responses actually occurred. Since the latency distribution of responses (See Figure 1) exhibited no apparent peaking outside the range from 1.0 to 3.5 seconds, only responses within these limits were included in the amplitude and magnitude analyses. Differences with respect to mean EDR latencies and mean EDR magnitudes were evaluated by an analysis of variance, a design designated by Lindquist (*6*) as Type III, expanded to include two levels of a fourth factor.

With respect to mean EDR magnitudes, differences among the five intensities and also between the two durations were significant at the one per cent level, as indicated by the analysis of variance results summarized in Table 1. Mean EDR magnitudes were greater for the short tone group than for the long tone group. The interaction of duration

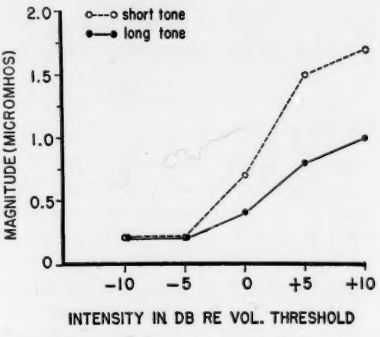

FIGURE 2. Relation between mean EDR magnitude and stimulus intensity.

with intensity, however, was also significant at the one per cent level. This interaction may be seen in Figure 2 which shows that the rate of the rise in mean EDR magnitudes with increased intensity (after —5 db) was greater for the short tone group than for the long tone group. Because of the significant interaction, the simple effect of intensity was tested separately for each of the two durations

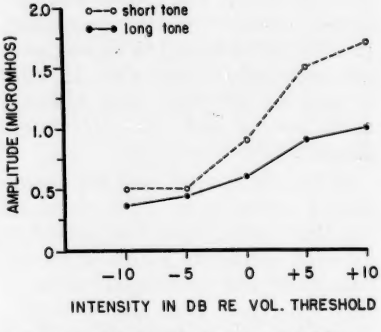

FIGURE 3. Relation between mean EDR amplitude and stimulus intensity.

TABLE 2. Analysis of variance of EDR latencies.

Source	*df*	*ms*	*F**	$F_{.01}$
Between subjects	(39)			
Duration (D)	1	27.46	2.30	
Schedule (S)	1	2.02	0.17	
Hand (H)	1	6.31	0.53	
D x S	1	31.59	2.65	
D x H	1	0.01		
H x S	1	9.28	0.78	
D x S x H	1	3.27	0.27	
Subjects in Procedure (Error Between)	32	11.93		
Within subjects	(160)			
Intensity (I)	4	182.20	9.83	3.47
D x I	4	3.53	0.19	
I x S	4	18.23	0.98	
I x H	4	1.72	0.09	
I x D x S	4	3.55	0.19	
I x D x H	4	4.53	0.24	
I x S x H	4	3.75	0.20	
I x D x S x H	4	0.67	0.04	
Subjects Within x Intensity (Error Within)	128	18.53		
Total	199			

*All between effects tested by mean square for error between; all within effects tested by mean square for error within.

of tone. Differences among intensities were significant[3] for each duration.

Although analysis of variance was not done on EDR 'amplitudes,' the relationship between EDR 'amplitudes' and tonal intensity, as shown in Figure 3, was found to be essentially similar to that between 'magnitudes' and tonal intensity.

Analysis of variance was also performed to evaluate differences with respect to mean EDR latency. For the latency analysis the total recorded range of response latencies from zero to 7.4 seconds was examined. The following criteria were used in selecting the responses included in this analysis: (1) if one or more responses occurred within the range from 1.0 to 3.5 seconds, the earliest one was chosen; (2) if no response occurred within the 1.0 to 3.5 second range, the earliest response within the overall 7.4 second interval was chosen. The latency distributions of responses meeting these criteria are shown in Figure 1. 'No responses' were assigned values of 8.0 to 22.0 seconds taken from a table of random numbers. The results of the analysis of variance, summarized in Table 2, indicate that differences

[3]For the short duration $F = 38.42$ and for the long duration $F = 22.96$. With 4 and 80 *df*, $F_{.01} = 3.56$.

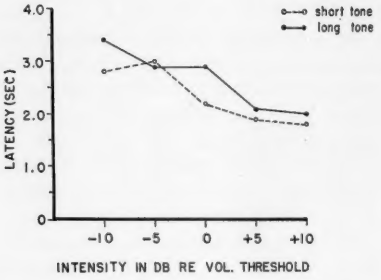

FIGURE 4. Relation between mean EDR latency and stimulus intensity.

among intensities were significant at the one per cent level of confidence. As may be seen in Figure 4, EDR latency was shorter with increased tonal intensity.

An examination of all of the EDR latencies at each of the five intensity levels indicates that the variability of latencies decreased markedly as tonal intensity increased from threshold to above-threshold intensities. This is revealed in Figure 1. This variability was less for the short tone group than the long tone group. A Bartlett test, performed on the EDR latencies at each of the five intensities, showed that there was significantly less variability of EDR latencies for the short tone group at 0 db, +5 db and +10 db *re* voluntary threshold, than for the long tone group.

Comparison of the resistance method results with the potential EDR's revealed several facts: (1) EDR latency for the potential method was essentially the same as that for the resistance method; (2) with the particular electrode placements used in this study, the wave shape of the potential response was more complicated than the resistance response; most of the time the potential wave was multiphasic; (3) within a particular subject's record, the potential response was less stable, showing changes in wave form throughout the test. Because of these variations in wave form, analysis of EDR potential amplitudes was not attempted.

Discussion

It is apparent from this study that changes in the intensity of audible tones significantly affect the dimensions of the EDR. When a tone was increased from −5 db to +10 db *re* voluntary threshold, EDR latencies became significantly shorter and the variability of these latencies smaller. EDR heights became significantly larger. Subjects exposed to the short tone method showed significantly less variability of EDR latencies than those exposed to the long tone method at these stimulus intensities.

Greater strength of conditioning, as evidenced by larger EDR heights, was achieved for the short tone group than for the long tone group. Moreover, subjects under the short tone method were more successfully conditioned than those under the long tone method. In order to obtain 40 conditionable subjects for the study, it was necessary to test 56 students; 10 of these were rejected due to difficulty in conditioning. Of these 10, nine failures came from the long tone group while only one came from the short tone group. Thus, for the long tone group, there was only 69% success in conditioning while for the short tone group there was 95% success in conditioning.

The finding that mean EDR latency and amplitude are orderly functions

of tonal intensity places EDR audiometrists closer to finding a method of rejecting spontaneous responses with a greater degree of certainty than has existed in the past. In addition, these data provide more information on the mechanism by which the EDR threshold is established, a consideration that will be treated in a succeeding paper (*4*).

Summary

EDR latency and amplitude as a function of tonal intensity were investigated in 40 subjects, 20 of whom were conditioned with a 0.6 second tone followed instantaneously by shock while the remaining 20 were conditioned with a 3.0 second tone followed instantaneously by shock. Analysis of EDR latencies and amplitudes showed that latencies and amplitudes were different for different tonal intensities.

The 0.6 second method resulted in higher strength of conditioning than the 3.0 second method as evidenced by larger EDR amplitudes and greater numbers of successfully conditioned subjects. The 0.6 second method resulted in a greater rate of change of amplitudes with tonal intensity than for the 3.0 second method. The variability of EDR latencies was smaller for the 0.6 second method than for the 3.0 second method.

References

1. Davis, R., Factors affecting the galvanic reflex. *Arch. Psychol.*, 18, 1930, No. 115, 1-64.
2. Goldstein, R., and Derbyshire, A., Suggestions for terms applied to electrophysiologic tests of hearing. *JSHD*, 22, 1957, 696-697.
3. Goldstein, R., Ludwig, H., and Naunton, R. F., Difficulty in conditioning galvanic skin responses: its possible significance in clinical audiometry. *Acta Otolaryng.*, 44, 1954, 67-77.
4. Hind, J. E., Aronson, A. E., and Irwin, J. V., GSR auditory threshold mechanisms: instrumentation, spontaneous response and threshold definition. *JSHR*, 1, 1958, 220-226.
5. Lehroff, I., An Experimental Study of Auditory Threshold Acuity of Cerebral Palsy Children by PGSR and Other Techniques. Ph.D. dissertation, University of Southern California, 1954.
6. Lindquist, E. F., *Design and Analysis of Experiments in Psychology and Education.* Boston: Houghton Mifflin, 1953.
7. Meritser, C. L., and Doerfler, L. G., The conditioned galvanic skin response under two modes of reinforcement. *JSHD*, 19, 1954, 350-359.
8. Schlosberg, H., and White, C. T., Degree of conditioning of the GSR as a function of the period of delay. *J. exp. Psychol.*, 43, 1952, 357-362.
9. Statten, P., and Wishart, D. E. S., Pure-tone audiometry in young children; psychogalvanic-skin-resistance and peep-show. *Ann. Otol., etc., St. Louis*, 65, 1956, 511-534.
10. Stewart, K. C., Some basic considerations in applying the GSR technique to the measurement of auditory sensitivity. *JSHD*, 19, 1954, 174-183.

GSR Auditory Threshold Mechanisms: Instrumentation, Spontaneous Response And Threshold Definition

J. E. Hind

A. E. Aronson

J. V. Irwin

Although the conditioned GSR or, as recently suggested (*2*), the EDR (electrodermal response) is now widely employed as an indicator of auditory sensitivity, the audiological literature reflects little interest in the precise manner in which the threshold occurs. Thus, the threshold is typically defined in terms of the percentage of responses to a series of stimuli in a manner analogous to the familiar finger-raising response of clinical audiometry. Although some workers have taken EDR amplitude into consideration to the extent of invoking an arbitrary minimum amplitude criterion for acceptance of a response (*6*), in general the role of decreasing amplitude in establishment of the threshold has not been assessed.

Another factor which has needed further attention has been the effect upon threshold of spontaneous activity and the problem of differentiating these non-auditory EDR's from true responses. Criteria intended to identify true responses on the basis of latency have been only partially successful. In one study it was found for the individual subjects that on approximately 20% of the silent control trials, responses occurred which could not be distinguished from tone responses on the basis of latency (*6*). The possibility that improved identification of spontaneous responses might be achieved through the use of a combined amplitude-latency criterion appears not to have been investigated.

Against this background, plans were formulated for the present research with the following objectives: (1) the development of EDR recording

J. E. Hind (Ph.D., Chicago, 1952) is Assistant Professor of Physiology in the Laboratory of Neurophysiology, University of Wisconsin; A. E. Aronson (Ph.D., Wisconsin, 1957) is Assistant Professor of Speech, University of Wisconsin; and J. V. Irwin (Ph.D., Wisconsin, 1947) is Professor of Speech and Director of the Speech and Hearing Clinic, University of Wisconsin. This study was aided by grants from the Alfred Laukhuff Trust for research in the field of cerebral palsy, from the National Institute of Neurological Diseases and Blindness (B-896) and from the Research Committee of the University of Wisconsin out of funds supplied by the Wisconsin Alumni Research Foundation.

equipment capable of meaningful amplitude measurement together with precise latency determination; (2) description of the behavior of mean response amplitude and latency in the vicinity of the absolute auditory threshold under two tone-shock intervals (0.6 and 3.0 seconds); (3) an attempt to discriminate spontaneous responses from true responses on the basis of amplitude and latency; (4) investigation of the effect of varying threshold criteria. The scope of this paper is restricted to objectives (1), (3) and (4). A description of the behavior of mean EDR amplitude and latency at near threshold intensities is published elsewhere (*1*).

EDR Instrumentation

For recording the EDR, a technique long used by neurophysiologists in the measurement of response latency was borrowed.[1] In this method the responses are displayed on a cathode ray oscilloscope with the beginning of the horizontal sweep of the beam synchronized with the stimulus; thus, position of responses on the scope face clearly defines latency. Visualization of responses against the noise background is facilitated by using a recording system of restricted bandwidth. The lower cut-off frequency is chosen just low enough for reasonable reproduction of what is considered to be the salient feature of the response wave form—namely, the onset. Use of a capacitor-coupled recording amplifier with an overall low-frequency time constant of one second results in fairly accurate reproduction of EDR onsets while eliminating slower fluctuations (*5*). An upper half-power frequency limit of 100 cps was chosen, because, although this is much higher than required for adequate reproduction of responses, recognition of the relatively sharp wave forms encountered in movement artifacts is made easier.

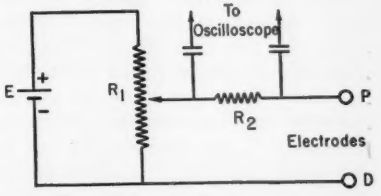

FIGURE 1. Circuit for measurement of EDR conductance changes. E = 10 volt battery, R_1 = 1000 ohm linear potentiometer, R_2 = 300 ohm resistor, P = palmar (positive) electrode, D = dorsal (negative) electrode.

This system provides for recording EDR's both in the form of internally generated potential differences and as changes in current flow from an external potential source. Only the method for the latter measurement will be described here on the assumption that response amplitude is best expressed as a change in electrical conductance, for this measure is relatively independent of the base value and tends to be normally distributed (*3*). Further, since the observed conductance varies with the magnitude of the measuring current, it is desirable that this magnitude be under experimental control.

The circuit, shown in Figure 1, is essentially that of Lacey and Siegel (*4*). With this arrangement the experimenter, by periodic adjustment of potentiometer R_1, maintains the base

[1]The authors are indebted to C. N. Woolsey, M.D., Professor of Neurophysiology, University of Wisconsin, for recommending the application of the cathode ray oscilloscope to this problem.

(between-response) current through each subject at the same constant value, in this case 40 μa. This base current can be measured most conveniently by a low resistance microammeter in series with the subject or, alternatively, the oscilloscope can be adapted to this purpose. The potential variations across R_2 are capacitor-coupled to the differential amplifier of the oscilloscope which has an adjustable full-screen sensitivity of from 0.2 to 20 mv.

By definition, the change in conductance is the ratio of the change in current to the change in applied voltage. Since the battery, R_1, and R_2 constitute a potential source of low internal resistance compared with the subject, the applied electrode voltage for any setting of R_1 remains essentially constant during the variations in current associated with a response. The magnitude of this applied potential may be read from the calibrated dial of R_1. By Ohm's law the transient potential variation across R_2 and the resulting scope deflection may be expressed in terms of the *change* in subject current. By dividing the measured change in subject current in microamperes by the applied electrode potential in volts the response amplitude is given directly in micromhos of conductance change.

Each sweep of the scope is photographed on a single frame of 35 mm recording paper so that responses may be analyzed for latency and amplitude. In this study an overall sweep duration of 7.4 seconds was employed.

Experimental Procedure

The experimental procedure is given in detail in another article (*1*) and hence will be outlined only briefly in this report. The subjects were 40 college students who were found to have normal voluntary thresholds on the right ear at 1000 cps. They were conditioned by pairing 40% shock reinforcement with a 1000 cps tone 10 db above each subject's voluntary threshold. After the subject gave seven responses to unreinforced tones, a random pattern of test tones was presented which included eight tones, five without shock, at each of five intensities: +10, +5, 0, −5, −10 db relative to each individual's voluntary threshold. The subjects were divided into two groups of 20 each, one receiving tones of 0.6 second in duration and the other receiving tones of 3.0 seconds in duration. Since the shock began immediately after the cessation of the tone, these tone durations can also be described as tone-shock intervals.

Results

Quite orderly relations were found between tone intensity and group mean EDR amplitudes and latencies. Also, significantly stronger conditioning occurred in the subjects exposed to tones of 0.6 second than in the subjects exposed to tones of 3.0 seconds. However, the objective of this paper is, first, to determine whether the spontaneous activity can be differentiated from the true responses on an amplitude-latency basis and, second, to observe the effects of varying the threshold criterion. Consideration here shall be limited to the 20 subjects receiving tones of 0.6 second although similar results occurred in the group receiving tones of 3.0 seconds.

In order to compare response amplitudes among the subjects, the conductance changes measured in each

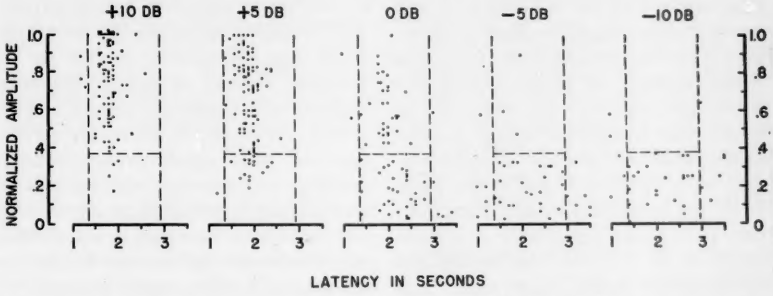

FIGURE 2. Relation between normalized EDR amplitude and latency for all responses of all subjects.

individual have been normalized by taking the largest response as 1.0 with all other responses expressed as a decimal fraction of each individual's largest response. Figure 2 shows the result of plotting the entire collection of normalized response amplitudes against latency at each intensity. There were 100 possible responses at each tone intensity since each of the 20 subjects received five unreinforced tones at each intensity. The plots show only responses having latencies between 1.0 and 3.5 seconds. These limits were chosen by inspection of the frequency distributions of response latencies which indicated that the range from 1.0 to 3.5 seconds would be broad enough to include essentially all of the true responses; to put it another way, this range includes that portion of each frequency distribution which is raised above the uniform background distribution.

It is clear from these plots that normalized amplitude drops off with decreasing tone intensity. The total number of responses also declines along the same direction. On the basis of the uniform distribution of response latencies and the low amplitudes, all of the responses at —10 db are interpreted as spontaneous or unrelated to the tone. This interpretation is strengthened by the observation of relatively little increase in response measures at —5 db compared with —10 db.

Both latency and amplitude criteria were applied to these data to determine how completely such criteria can be employed to reject the responses at — 10 db without decreasing appreciably the number of acceptable responses at +10 db. The latency

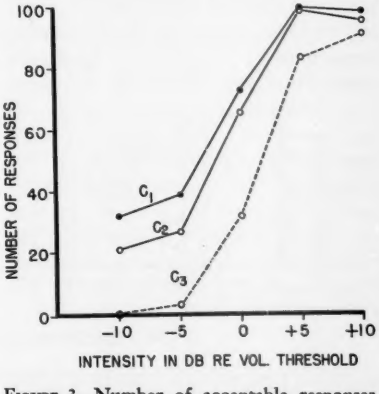

FIGURE 3. Number of acceptable responses at each intensity using the three amplitude-latency criteria described in the text.

limits of 1.4 to 2.9 seconds indicated by the two dashed vertical lines on each plot were chosen as being just broad enough to avoid rejecting a significant number of responses at those tone intensities which yielded clear peaks of latency distribution, that is, at +10, +5, and 0 db. Similarly, the horizontal dashed lines indicate a minimum acceptable amplitude criterion of 0.38 on the normalized conductance scale. This value was selected to reject 100% of the responses at −10 db which fell within the 1.4 to 2.9 second latency criterion.

The effects of these arbitrary amplitude-latency criteria on the number of acceptable responses at each intensity are summarized in Figure 3. Here, C-1 refers to the most lenient criterion: namely, responses of all amplitudes and latencies from 1.0 to 3.5 seconds. This is, of course, the total collection of responses shown in Figure 2. At −10 db there are 32 responses; thus spontaneous responses meeting the latency criterion of 1.0 to 3.5 seconds occurred on about 30% of the below-threshold trials. When all amplitudes are accepted but the acceptable latency range is narrowed to 1.4 to 2.9 seconds, criterion C-2 results. This drops the number of acceptable responses at −10 db to 21, a decrease of 34% in spontaneous activity with a decline of only four responses or approximately 4% at +10 db. Finally, when the tightened

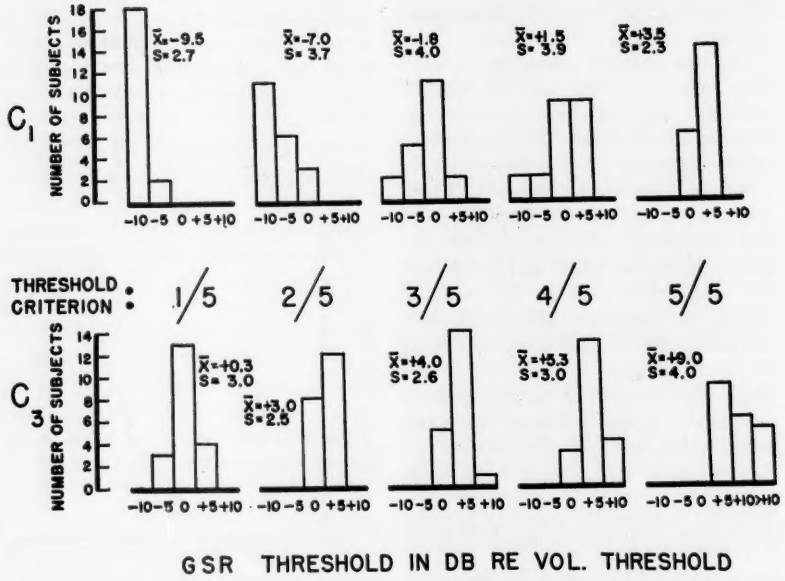

FIGURE 4. Frequency distributions of EDR thresholds for 20 subjects based on five threshold criteria. C-1 thresholds determined by using all responses with latencies of 1.0 to 3.5 seconds regardless of amplitude; C-3 thresholds determined by using only responses of 0.38 normalized amplitude or greater and latencies from 1.4 to 2.9 seconds.

latency criterion is combined with the 0.38 or greater amplitude requirement, criterion C-3 results. Here, by intention, all responses are eliminated at −10 db while only seven responses (about 7%) are rejected at +10 db.

The results obtained by the application of five definitions of EDR to the same set of individual data are shown in Figure 4. This figure displays the two sets of threshold distributions obtained with the two criteria of acceptable response, C-1 and C-3. Each distribution includes 20 threshold values, one for each subject. The EDR thresholds are expressed in decibels relative to the voluntary threshold. In addition, the chart illustrates the effect of varying the number of responses required for definition of threshold from one out of five through five out of five. Given on each distribution are the mean threshold, $\overline{X}$, and the standard deviation, S.

Since all threshold definitions are arbitrary, it is appropriate to ask which of these threshold criteria is most desirable. One can be eliminated rather easily, the one out of five criterion for C-1 responses, for this yields a distribution in which 18 of the 20 subjects have an EDR threshold of −10 db; this results from the fact that with a broad latency criterion and no amplitude criterion most subjects have at least one out of five spontaneous responses. Thus the apparent threshold would occur at the lowest intensity presented, which in this case was −10 db. As the number of responses required for threshold definition is increased, the mean threshold also increases.

The best EDR threshold criterion could be interpreted as the one which gives best correlation with the voluntary threshold. On these plots that would correspond to a mean EDR threshold of zero db. However, it must be remembered that perfect agreement between the two thresholds need not be anticipated under the present experimental conditions, for the EDR and voluntary thresholds were obtained at different times and with different temporal patterns of tonal intensities. One other means of judging the adequacy of threshold criteria is the spread or standard deviation of the collection of thresholds. For the C-1 responses the standard deviation reaches a minimum of 2.3 db at the five out of five criterion, which gives a mean EDR threshold of +3.5 db. For the C-3 or decontaminated responses, a minimum standard deviation of 2.5 db is reached at a criterion of two out of five responses with a mean EDR threshold of +3.0 db, a value in good agreement with the C-1 five out of five criterion.

Discussion

These data demonstrate that spontaneous activity can be differentiated from true responses, at least under these experimental conditions. From the standpoint of practical use of EDR threshold tests, however, it would appear to be more expedient to raise the required threshold criterion to a value between 80% to 100% response rather than to attempt to eliminate spontaneous activity by amplitude-latency criteria. If the usual audiometric threshold criterion of 50% responses is used and spontaneous activity is not rejected, the EDR threshold can clearly appear to be lower than the voluntary.

The particular amplitude-latency criteria employed in this study will probably not be appropriate for other

experimental situations. As Zuidema *et al.* (7) have pointed out, the amplitude of specific or evoked responses and the amount of non-specific or spontaneous EDR activity both vary along a continuum which parallels the state of alertness of the subject. This ranges from complete lack of alertness in sleep through maximum attentiveness to wild panic. Successful use of the conditioned EDR for auditory threshold determination requires that the operator maintain the subject in the range near the center of the continuum where the amplitude of specific response is at a maximum and the number of non-specific responses is not prohibitively high.

Summary

EDR recording apparatus which fulfills the objectives of meaningful amplitude measurement together with precise latency determination is described. Amplitude is measured as the change in electrical conductance of the skin using a constant base (between-response) current of 40 μa. Latency determination is facilitated by presenting the EDR's on a cathode ray oscilloscope, each sweep being triggered simultaneously with the stimulus tone onset.

Data from 20 subjects are analyzed by plotting normalized amplitude *vs.* latency at each of five test-tone intensities: +10, +5, 0, −5, and −10 db *re* each subject's voluntary threshold.

All of the responses at −10 db are interpreted as spontaneous, that is, unrelated to tone. An arbitrary amplitude-latency criterion is chosen to reject all of these responses at −10 db; when this criterion is applied to the +10 db intensity only seven per cent of the responses are rejected.

The effect of varying the threshold definition from one out of five responses to five out of five is illustrated both for the total collection of responses and the collection remaining after rejecting those responses which fail to meet the amplitude-latency criterion. For the total collection of responses, which meet the usual clinical definition of response, least spread of the set of 20 thresholds occurs with a five out of five criterion. The results indicate the advisability of using a high percentage response for a threshold criterion in the presence of spontaneous activity.

References

1. Aronson, A. E., Hind, J. E., and Irwin, J. V., GSR auditory threshold mechanisms: The effect of tonal intensity on amplitude and latency under two tone-shock intervals, *JSHR*, 1, 1958, 211-219.
2. Goldstein, R., and Derbyshire, A. J., Suggestions for terms applied to electrophysiologic tests of hearing. *JSHD*, 12, 1957, 696-697.
3. Lacey, O. L., and Siegel, P. S., An analysis of the unit of measurement of the galvanic skin response. *J. exp. Psychol.*, 39, 1949, 122-127.
4. Lacey, O. L., and Siegel, P. S., An improved potentiometric circuit for measuring the galvanic skin response. *Amer. J. Psychol.*, 61, 1948, 272-274.
5. Stewart, K. C., A new instrument for detecting the galvanic skin response. *JSHD*, 19, 1954, 169-173.
6. Stewart, K. C., Some basic considerations in applying the GSR technique to the measurement of auditory sensitivity. *JSHD*, 19, 1954, 174-183.
7. Zuidema, G. D., Cohen, S. I., Silverman, A. J., and Riley, M. B., Human tolerance to prolonged acceleration. *Wright Air Develop. Center Tech. Rep. 56-406*, U. S. Air Force, Wright-Patterson Air Force Base, Ohio, Oct., 1956.

Myoelastic-Aerodynamic Theory Of Voice Production

Janwillem van den Berg

The myoelastic-aerodynamic theory of voice production deals with the control of the larynx by the higher centers of phonation. Only a few of the problems involved in the theory can be discussed here; therefore this paper will be restricted to normal voice production and the vowels.

In normal voice production the vocal organs are (1) the lungs and trachea, which act as a bellows and windpipe, (2) the larynx, a complex organ which forms the generator for the vowels and (3) the vocal cavities, which form the resonator system driven by the glottis-generator in the larynx. These organs make a unit, and it is somewhat artificial to separate them, but it is not possible to understand the function of the unit without a knowledge of the properties of the components. Therefore, it is reasonable to start with a brief discussion of the behavior of the different organs and to end with a consideration of the peculiarities evoked by their coupling together. Apparently, the larynx is the central organ for the vowels, and thus it is necessary to understand how the larynx is affected by the vocal cavities and the trachea and lungs. The effects are known when the acoustical resistance or impedance Z (Z=pressure/volume velocity) is known in the upward direction at the level just above the larynx and in the downward direction at the level just below the larynx.

The acoustical impedance of the vocal cavities at the level just above the larynx is directly related to the transmission of the sound signal from the larynx to the external air. This transmission was measured (*3*) with a hemilaryngectomized subject for 11 cardinal vowels of the International Phonetic Alphabet, as seen in Figure 1. The transmission T (T=volume velocity at the lips/volume velocity at the input of the vocal cavities) appears to depend upon the frequency. T is large for the formant or resonant frequencies of the vowel and is small for frequencies between the formants. Naturally the resonances of the vocal cavities are not infinitely sharp. In fact, the formants are damped (*4*) by energy losses resulting from (1) radiation of energy from the lips to the external air, particularly for the high formants, (2) friction in narrow parts along the vocal tract, particularly for formants of intermediate frequency and (3) sympa-

Janwillem van den Berg (D.Sc., 1953, Groningen University, The Netherlands), is a Lector in medical physics and in charge of the medical physics department of the physiological laboratory of Groningen University. This article is based on a paper read at the Chicago International Voice Conference, May 1957.

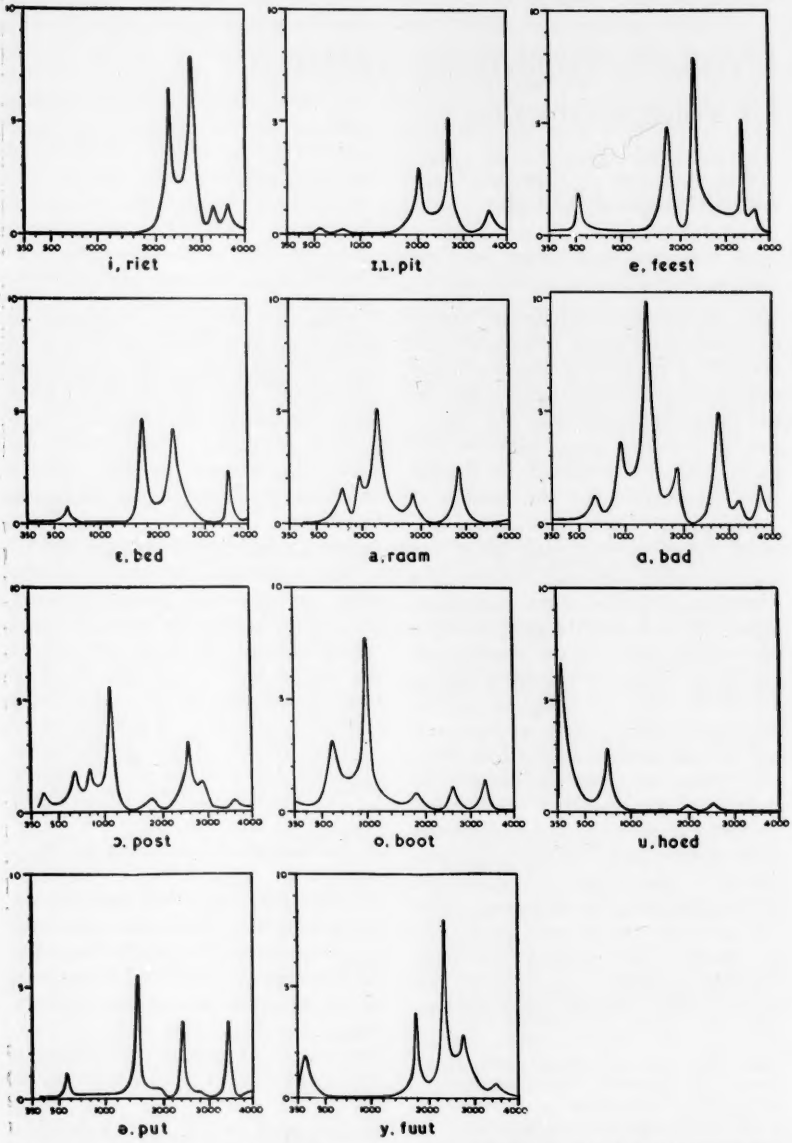

FIGURE 1. Transmission of the vocal cavities for 11 cardinal vowels of the International Phonetic Alphabet.

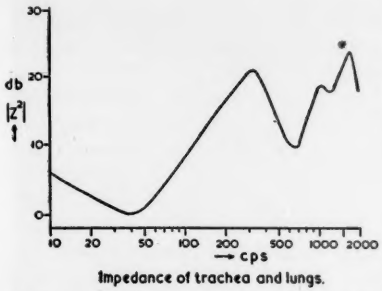

FIGURE 2. Input impedance of human trachea and lungs.

thetic vibrations of the walls of the cavities, particularly those of the throat in response to the low formants. With an open glottis, formants also lose vibrational energy towards the glottis, trachea and lungs. The acoustical impedance at the input of the vocal cavities resembles the transmission curve; the input impedance of the vocal cavities is very large at the frequencies of the formants, and a large supraglottal pressure is needed to vibrate the air. Dunn (*21*) introduced a very successful method for the calculation of the properties of the vocal cavities by representing them in an electrical analogue and applying the formulae for electrical transmission lines. The method also proved very successful for the design of electrical speaking machines (*21, 48*).

Interoceptive reflexes can arise at the palatum and the posterior wall of the pharynx; the sensory nerves involved are part of the fifth nerve. The magnitude of the reflexes depends on the vowel in question, on the sound level and on the acoustical properties of the room. The reflexes are primarily evoked by the second formant of the vowel (*2*). As the ear is the dominating proprioceptive organ in phonation, the effect of the above reflexes on the function of the higher centers is hard to investigate and not yet clear. The interoceptive reflexes in the larynx are conveyed by the tenth and eleventh nerve. These reflexes are primarily due to the first formant of the vowel (*2*).

It is virtually impossible to measure the acoustical impedance of the trachea and lungs with living subjects. Recent experiments (*5*) on dogs and human corpses yielded results like those of Figure 2. The first valley belongs to the resonance of the trachea and lungs as a whole; the following peaks and valleys are caused by subsequent resonances and antiresonances of the trachea and connected bronchi, which respond as an open organ pipe, with an effective length which decreases somewhat with increasing frequency. As in the vocal cavities, the acoustical impedance is large at the peaks of the curve. In measurements (*5*) of the amplitude

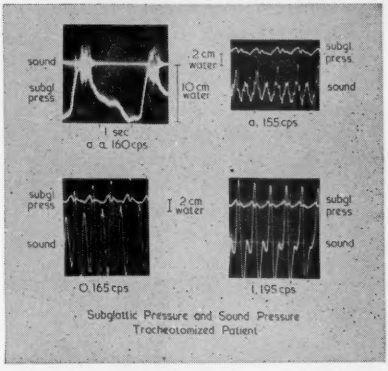

FIGURE 3. Sound pressure and subglottic pressure. Simultaneous recordings on a tracheotomized subject with a normal larynx. The subglottic pressures during vocalization amounted to about 10 cm water pressure. The variations during the vocal period may be compared with the 2 cm water pressure bar.

of the vibrations of the chest on living subjects singing the same vowel with constant volume velocity and varying pitch, f, it was found that the amplitude was proportional to f^{-3}. Experiments (*6*) on the sound pressure and the subglottic pressure with a tracheotomized subject (See Figure 3) were in agreement with these observations. In spite of their typical resonant behavior the trachea and lungs are very efficient as a bellows and windpipe.

Interoceptive reflexes in the trachea-lungs system, conveyed by the tenth nerve, may be active during loud speech or singing. It is, however, very difficult to investigate them. Presumably, these reflexes play a part in the so-called 'Atemstütze' (*36*).

The role of the larynx, the central voice organ of the vowels, has up to now been widely discussed. In ancient times, Galen thought the trachea to be the central organ, acting as a flute. Dodart (*18*, *19*, *20*) proved in 1700 that the vowel originates in the larynx. However, by analogy to mouth-whistling, in which he was an artist, he thought that the pitch depended upon the area of the glottis. In his opinion, the only task of the vocal folds and laryngeal muscles was to vary the area of the glottis appropriately; he believed that eddies in the glottis generated the sound, as with the whispered voice. In the nineteenth century the physicists Savart (*43*) and Guillemin (*28*) defended a comparable analogy. The vowel was supposed to originate by eddies in the laryngeal ventricle, a mechanism comparable to that in the hunter's whistle, which Savart attempted to explain. By experiments with the inhalation of gas mixtures with a high sound velocity, this theory was easily disproved (*9*), as the pitch remained constant, whereas the formants shifted in frequency according to the variation of the velocity of sound.

A very important contribution is due to Ferrein (*23*), 1741, who first performed animal experiments on voice production and proved that the vibrations of the vocal folds are essential and not of a secondary nature. He thought, however, that the sound was directly produced by the vibrations of the folds, in the manner that a string drives the surrounding air. Helmholtz (*29*) showed that the puffs or rushes of air escaping through the glottis are the primary source of the sound. The larynx thus contains a glottis-generator.

The essential function of the vibrations of the folds, at least with the chest, mid and falsetto voice, is now commonly accepted. The divergence between the present day theories arises when it is asked how these vibrations are produced. According to Husson's (*31*, *32*, *33*, *34*) neurochronaxic theory, each new virbation is induced by a new impulse sent by the brain to the musculus vocalis by way of the recurrent nerve. The myoelastic-aerodynamic theory, on the other hand, postulates that the vocal folds are actuated by the stream of air delivered by the lungs and trachea. Thus, the fundamental frequency of the vibration is thought to depend on the effective mass and stiffness of the vocal folds; they are regulated, in all fine detail, by the sustained innervation of internal and external laryngeal muscles and by the activity of the bellows and the associated resonators. These theories cannot be united and anatomical, histological, physiological and physical data will have to provide the decision. These data are available to some extent, but they are differently

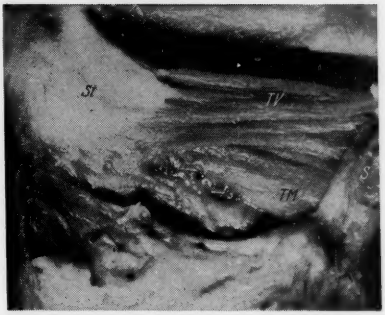

FIGURE 4. Left musculus vocalis: mucous membrane, ligamentum vocale and intermuscular connective tissue removed (S = thyroid, St. = arytenoid, TV = portio thyreovocalis and TM = portio thyreomuscularis).

interpreted.

Anatomically, the larynx is one of the most complicated structures of the body. A great number of muscles attach to the arytenoids; the anatomy of the musculus thyreoarytenoideus internus, the musculus vocalis, has been much discussed since the investigations of Goerttler (27) and Behringer (2). Their studies seemed to provide an anatomical substantiation for the neuro-chronaxic theory. They concluded that the musculus vocalis is made up of two bundles, each inserted on the border of the vocal folds, the ligamentum vocale. With such an insertion it would be evident that a contraction of the musculus vocalis would open the glottis, as required by the neuro-chronaxic theory. As has been pointed out (*9, 11*), this anatomy would be highly insufficient for the vital function of the larynx to close the trachea and lungs hermetically from the external air (which can also be done by species without ventricular bands, that is, ruminants). For example, during a cough or during swallowing, the larynx would open at a contraction of the musculus vocalis. Investigations by Wustrow (*58*), Mayet (*37, 38*) van den Berg and Moll (*10*), and Vosteen (*54*) yielded results in agreement with the classical opinion, showing a musculus vocalis with longitudinal fibres, inserting on the thyroid and on the arytenoid. The musculus vocalis appeared to be made up of two bundles (See Figure 4) which Wustrow called Portio thyrovocalis and Portio thyromuscularis according to their insertion on the processus vocalis or muscularis of the arytenoid. Only a negligible number of muscle fibers were found to insert on the ligamentum vocale, whereas fibers do insert on the free border in the region of the conus elasticus. Presumably, Goerttler's conclusions (27) were based on a wrong interpretation of the twisting of these bundles with respect to each other and of the insertion of muscle fibers on the conus elasticus.

The innervation of the larynx is very rich. The motor supply of the internal muscles comes from the recurrent nerves, which also convey the proprioceptive reflexes and possibly also the interoceptive reflexes of the lower part of the larynx. Interoceptive reflexes of the mucous membranes of the upper part of the larynx are conveyed by the superior laryngeal nerve, which also contains the motor and sensory nerves of the musculus cricothyreoideus. The muscle fibers of the musculus vocalis and the nerve fibers of the recurrents are of the rapid type and a great number of end organs in the musculus vocalis suggest numerous sources of proprioceptive reflexes. All available data indicate that the larynx is well equipped for the tremendous task it must perform daily.

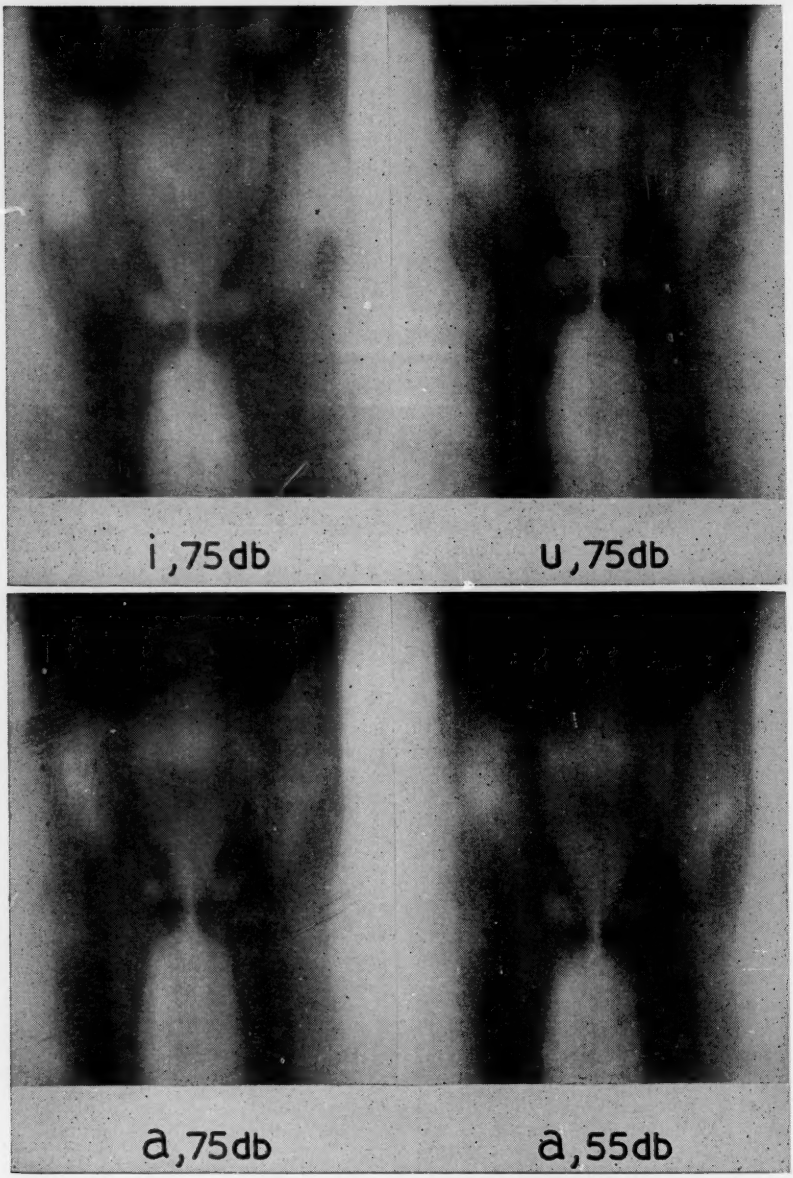

FIGURE 5. X-ray tomograms, 2 cm behind Adam's apple, of a bass vocalist, chest voice at F, 173 cps, distance 1 m.

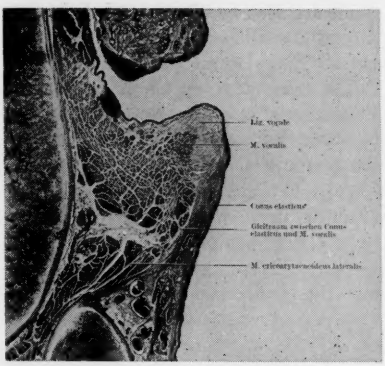

FIGURE 6. Frontal coupe through the vocal fold of an adult. Haematin-Eosin (29).

The internal laryngeal muscles form a structure which is entirely different from that composed of two elastic bands which may be brought together by the action of the musculus interarytenoideus. On the contrary, the form of the vocal folds can be varied and this fact needs to be taken into account (*9, 11*). In the falsetto voice the margins are seen to be rather thin and pointed, and the vibrations are limited to a small region in the neighborhood of the ligamentum vocale; this shape results from the high tension exerted upon the ligamentum by the contraction of the musculus cricothyroideus, which lengthens and stretches the vocal folds. The investigations showed that the ligamentum vocale needs large forces in order to be stretched some 20 to 40%, and this is one explanation for the small amplitude of the vibrations in the high pitches of the falsetto voice. The chest voice, on the other hand, shows a thicker margin of the folds in which case they can make contact over a depth of several millimeters. This depth is related to the sound level (See Figure 5). With an intense sound, the cords may be firmly pressed together (*11*) by the contraction of the musculus vocalis externus and the musculus cricoarytenoideus lateralis. With the chest voice the tension mainly depends on the contraction of the musculus vocalis and a greater part of the folds is put into vibration than with the falsetto voice. The amplitude of the vibrations and the effectively vibrating mass of the folds in chest voice is thus much larger than with the falsetto voice. The neglect to consider this ability of the larynx was one of the reasons for the introduction of the neurochronaxic theory (*9, 11*).

In discussions on the ligamentum vocale (See Figure 6), it is mostly tacitly assumed that this ligament is an elastic string of negligible thickness. In fact, however, it is a thin band, which can be three to five millimeters broad. The cells of its mucous membrane are modified into plate epithelial cells, which are usually found in places where relatively high contact pressures can occur. Presumably, this modification is caused by the contact of the vocal folds during the chest voice. The upper and lower edges of the ligamentum can be clearly seen on alternately preparing for a low pitch tone in the expiratory and inspiratory voice. Thus, the glottis, which we define as the region between both ligaments, has a certain depth, which depends on the type of the voice and the sex. In the low pitch chest voice the longitudinal tension of the vocal folds arises from the musculus vocalis; in the high pitch falsetto voice it arises from the ligamentum vocale.

The thickness of the margins in the chest voice causes a phase difference between the vibrations in different horizontal planes of the larynx. To

cite the description issued with the Bell film[1] of 1940: 'At the lowest pitches . . . vibration takes place throughout a considerable depth, from above down towards the trachea. Assuming the closed position of the cycle the opening appears to begin from underneath, toward the trachea, progressing upward and outward. In other words, there exists a phase difference between different vertical positions.' According to the photographs the glottis opens from beneath to above and closes in the same direction, at least with the expiratory voice. This phase difference is due to the fact that the ligamentum vocale (See Figure 6) is not stretched in the low pitches of the chest voice, permitting large deformations and large amplitudes of the vocal folds. These large deformations are further possible because the fibers of the musculus vocalis can shift somewhat with respect to each other and because of the looseness of the upper surface of the vocal folds. These factors explain the occurrence of rapid wavelike motion or ripples which are seen to pass over the top surface after the opening and closing of the glottis. As shown by the Bell film and later films, 'Horizontally the opening along the length of the cords may also have a phase difference.' This is due to an unequal distribution of mass and stiffness along the cords; in the lowest pitches of the chest voice the arytenoids are seen to vibrate and their specific mass differs from that of the cords. The lateral pressure can also differ along the length of the cords. Phenomena arising from differences between left and right cords will be discussed in a following paragraph.

With high-speed photography or modern synchronous stroboscopy, light reflections from the curved top surfaces are sometimes going inward; but the above wavelike motions or ripples over the top surface of the cords, after the opening of the glottis, are always seen to pass outward from the glottis to the lateral side. The reversed direction will be seen with non-synchronized stroboscopes when the pitch of the subject is somewhat below the frequency of the stroboscope instead of above. Presumably this is the explanation for the results of Heymann (*30*), who, with a non-synchronized stroboscope, in 80 per cent of his normal cases in the chest voice observed an 'Art von aktiver wellenförmiger Bewegung beider Stimmlippen . . ., die sich beiderseits vom Sinus Morgagni her nach dem freien Rand der Stimmlippen hin bewegt und dort verschwindet bzw. von der passiven, durch den Luftstrom bedingten Schwingung des freien Randes übernommen wird.' He thought that this was caused by an active innervation: 'Jedenfalls sieht man das Phänomen nie bei einseitiger Nervenschädigung, wie zum Beispiel einer Rekurrensparese, auf die erkrankten Seite.' This explanation is the more likely, as Heymann did not discuss this source of error. The disappearance of the ripples on the top surface on the paretic side will be caused by the much smaller amplitude of the vibrations at this side, and eventually also by a decrease of the water content of the mucous mem-

[1]High-speed motion pictures of the human vocal cords have been made for Bell Telephone Company, 1940, by J. C. Steinberg, D. W. Farnsworth, and H. J. Smith (available at the Bureau of Publications, Bell Telephone Company). Later high-speed films of the human larynx have been prepared by R. Luchsinger, B. Vallencien, P. Moore and H. von Leden. A recent high-speed film by J. Piquet, G. Decroix and J. Libersa will be discussed later.

brane at this side.

In his rubber models of the larynx, Smith (*44*) assumes a tight contact between the musculus vocalis and the ligamentum vocale, but totally free movability of the mucous membrane on the conus elasticus with respect to the underlying muscle. In these models, the vertical phase differences in the chest voice are supposed to be due to the free movements of these subglottic membranes. There is, however, no anatomical basis for these models which cannot represent the wavelike motions passing over the top surface of the folds. The edges of the models are relatively stiff, and the 'glottis' must be made too deep in order to show the phase differences. The role of the mucous membranes is overestimated in these models.

The role of the musculus vocalis is overestimated in the neuro-chronaxic theory (*31, 32, 33, 34*), which supposes that each vibration is started by a contraction of the musculus vocalis, opening the glottis. It was just seen that there is no anatomical basis for this theory and that such an anatomical structure would interfere with other vital functions of the larynx. There are, however, many more arguments to reject this theory. Some arguments may be briefly discussed now: (1) The left recurrent nerve is about 10 cm longer than the right one. At a velocity of 100 m/sec of the impulses along the nerves (the highest velocity observed in nerves), the impulse of the left nerve would arrive in the larynx with a delay of 1 msec with respect to the impulse from the right nerve. This is a short delay, but it is not negligible; for example, it amounts to half the duration of the cycle at a fundamental frequency of 500 cycles per sec. According to the theory, at such large frequencies large phase differences between the vibrations of the left and the right vocal folds ought to result; for example, one fold might be maximal outward and the other one maximal inward at the same time. (These phase differences, however, are only seen in pathological cases.) Theoretically, it might be possible to correct for such a time delay, but there are no adequate tools for this correction, as problems of the same kind arise with the propagation of the proprioceptive impulses back to the brain. The time involved in the whole chain, motor and sensory, is at least 5 msec, which is long compared with the duration of the vocal period at pitches of 300 cps and higher. It is interesting to note how the simultaneity of movements, for example, of the two hands breaks down when the period of the movements decreases and attains a value comparable to the duration of the propagation of the impulses along the nerves. (2) The peculiarities of the coupling of the vocal cavities to the larynx and the vocal folds to each other cannot be plausibly explained. (3) The theory cannot explain why a certain minimal air flow is needed to produce vibrations of the folds. The vibrations can be produced in the normal way, expiring the air, and also in the reversed way, inspiring the air, but there is a gap between these modes of vibration; the vibrations cease when the flow is too small. (4) The pitch does not alter on alternately expiring and inspiring air in the falsetto or mid voice, but does so in the chest voice. (5) The effect of the subglottic pressure on the pitch cannot be plausibly explained. (6) The effect of the resistance of the vocal cavities on the pitch cannot be plausibly explained. (7) The very high pitches attained by

some subjects—up to more than 2000 cps—are due to eddies in the larynx and subsequent cavity resonance, but not to vibrations of the vocal folds, as recently shown by experiments (*5*) after the inhalation of different gas mixtures. In fact, the highest pitch attainable was shifted in several experiments from 2250 cps in air to frequencies up to 3370 cps in the mixtures. (8) Active vibrations at a high pitch, that is, 500 cycles per sec, would tend to damage the folds in a short time, as the heat of activation of the muscle fibers is rather high, presumably of the order of 0.5 cal/gram/sec at a pitch of 500 cps. (9) Gutzmann's pressure test cannot be plausibly explained. (10) High-speed motion pictures of the vocal cords during a 'breathy attack' show that the vibrations start with an inward movement of the folds instead of an outward movement, as observed by Smith (*44*). (11) Sonninen (*46*) observed that the highest pitch attainable in the falsetto voice depends on the position of the head; the pitch drops when the head is bent backwards.

A recent high-speed film by Piquet, Decroix and Libersa (*40*) is claimed to show vibrations of the vocal folds without any stream of air through the glottis. The experimenters put a cannula in the trachea of a patient, via a tracheostoma, and inflated the cuff around the cannula in order to prevent any flow through the glottis. Then the patient was asked to phonate and the cords were photographed. Vibrations, with movements about equal to those during normal phonation, were recorded during about 0.5 sec. This would imply that the vibrations seen during normal phonation are also due to repetitive nerve impulses and not to actuation by the stream of air. However, it is probable (*12*) that the occlusion of the trachea was not adequate, as the observers heard a sound during the experiment, with a duration comparable to the duration of the film. A very small stream of air through the glottis can account for the observed vibrations; recently 9 to 10 ml of air were shown (*12*) to be sufficient for the production of a sound during 0.5 sec. Stroboscopic experiments (*12*), on a tracheotomized subject with a normal larynx showed that the sound and the vibrations of the folds cease immediately when the stoma is opened during a sustained effort to phonate; the sound and vibration resume immediately when the stoma is occluded again. When the stoma is alternately opened and closed, this sequence can be seen a number of times during one sustained effort to phonate. This phenomenon has been confirmed by von Leden and Moore with a number of high-speed films on a patient with a normal larynx: 'Repeated motion pictures on different days show with-

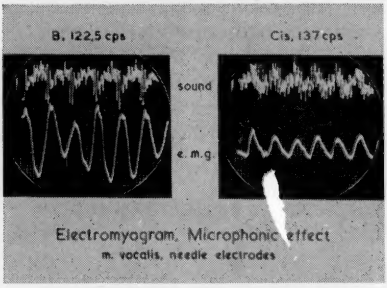

FIGURE 7. Microphonic effect of the larynx, vowel [a], fundamental frequency B, 122.5 cps, and C, 137 cps. Upper curves are sound pressure. Lower curves are electromyogram of a vocal fold, needle electrodes. The corresponding curves are synchronous. The passive variations (deformation, microphonic effect) are much greater than the active variations (spikes).

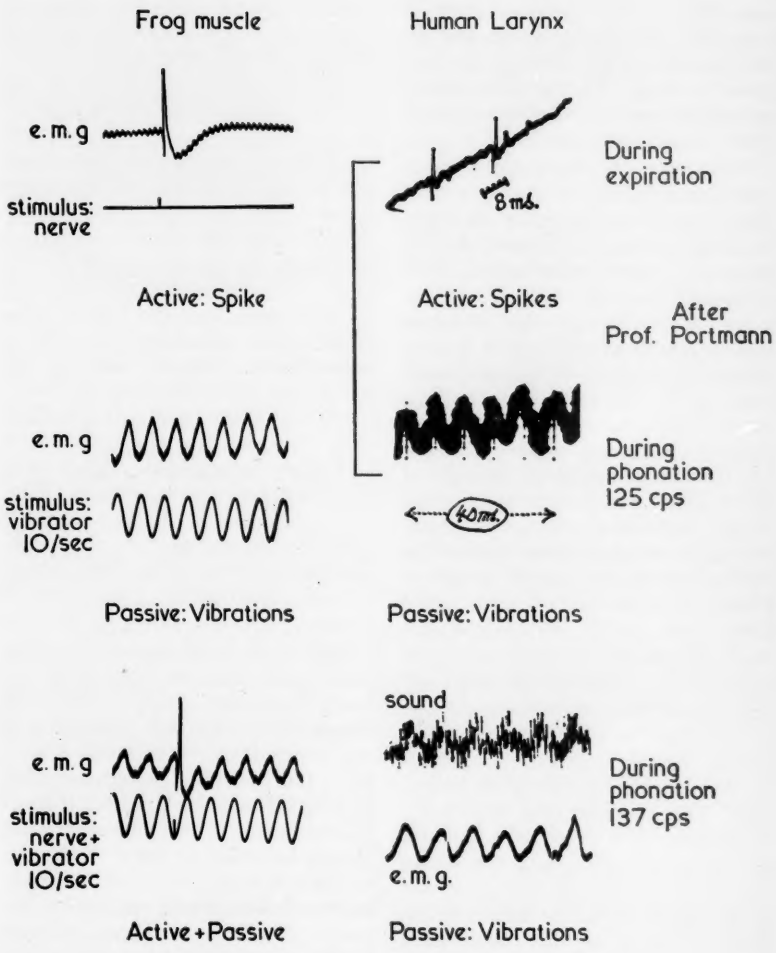

FIGURE 8. Electromyograms of a frog muscle and human vocal muscles, showing the difference between active and passive variations. Top curves are active variations (spikes) initiated by a nerve impulse. Middle curves are passive variations, caused by vibrations of the muscle on account of a vibrator actuating the frog muscle or a stream of air actuating the vocal muscle. Lower curves are frog muscle stimulated by a nerve impulse and a vibrator; vocal muscle is actuated by a stream of air, together with the sound emitted from the mouth; no initiating active variations are seen.

out doubt the complete cessation of vibratory movements when the air escapes unimpeded through the tracheal opening; the normal vibratory pattern resumes as soon as the tracheal stoma is again occluded.'[2]

Experiments by Moulonguet *et al.* (*39*) were claimed to show synchronism, during some parts of the registrations, between the impulses derived from a recurrent nerve and the fundamental frequency of the sound curve. Unfortunately, the reproduction of the curves in the article is not good enough to allow exact measurements by the reader and thus it is, for example, impossible to examine whether a number of the recorded impulses were sensory instead of motor. Moreover, the neuro-chronaxic theory requires synchronism during the *whole* sequence. Some parts of the curves suggest that a kind of microphonic effect may have been recorded too.

A microphonic effect of the larynx has been recorded by van den Berg and Spoor (*13*) during experiments on a patient with a stoma above the larynx, with needle electrodes put into the vibrating part of a normal vocal fold. The records (See Figure 7) show perfect synchronism between the electromyogram and the sound curve, but the e.m.g. is passive and is not caused by active muscle spikes in the rhythm of the fundamental frequency. This passive microphonic effect of the larynx, which is due to vibrations of the muscle fibers and the needle electrodes, can also be used to explain the results of Portmann (*41, 42*). When electromyograms are obtained under different conditions (with a resting muscle stimulated *via* its nerve, with a muscle vibrated passively by means of a vibrator put against the muscle or by a stream of air from the lungs, and with the combination of stimulation *via* the nerve and passive vibration) it is seen that active variations (spikes) are very rapid, with a duration of 1 msec or less, whereas passive variations are much slower (See Figure 7). The variation must have the typical form of the spike, otherwise it cannot be explained as a sign of muscular activity (See Figure 8). Recent experiments by Spoor and van Dishoeck (*47*) have shown a superposition of a passive microphonic effect and asynchronous muscle spikes. With electrodes placed in the non or slightly vibrating part of the cords, Faaborg-Andersen (*22*), have shown that the microphonic effect can be avoided.

Laget's (*35*) experiments on dogs were interpreted as supporting the neuro-chronaxic theory. With no stream of air through the glottis a recurrent nerve was stimulated with volleys of electrical impulses during ⅓ sec, followed by a resting time of ⅔ sec, etc. This procedure was chosen in order to prevent the development of a tetanus of the musculus vocalis, which normally develops at a rather low rate of continuous stimulation, 50 to 100 stimuli per second. Stroboscopic examination showed slow vibrations up to a stimulating frequency of 600 per second. At the arrival of the volleys, however, the fold was seen to move inward, but not outward as required by the neuro-chronaxic theory. Fessard and Vallancien (*24, 25, 26*) showed that the vocal cords are put into slow oscillations, with a frequency from two to

[2]This quotation is from a personal letter from Dr. von Leden of Chicago, on the subject of this film.

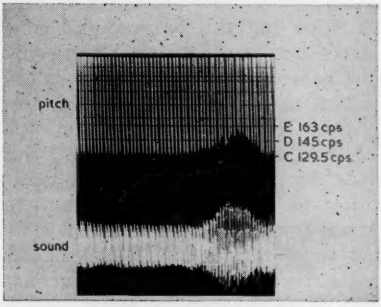

FIGURE 9. Effect of the subglottic pressure on the pitch. A sudden pressure is exerted on the stomach during a sustained vowel.

seven per second, when stimulated with high frequency impulses. With stroboscopic illumination, they argued, these slow oscillations are seen in their own, slow period. High-speed photography would reveal the true frequency, but such films have not been made. Moreover, with volleys of very short duration the phenomenon to be investigated is superimposed upon the large movements induced by the simultaneous stimulation of the motoric nerve fibers of the musculus interarytenoideus. This large movement will induce a damped oscillation of the musculus vocalis which must be separated from the movements caused by the stimulation of the vocalis fibers.

It may now be asked how the functioning of the glottis-generator is to be realized on the basis of the myoelastic-aerodynamic theory. The fundamental frequency of this generator is equal to the frequency of the vocal fold vibrations and this frequency depends on a number of interdependent factors: (1) the effective mass of the vibrating part of the vocal folds; (2) the effective tension in the vibrating part of the vocal folds; (3) the effective area of the glottis during the cycle, which determines the effective resistance of the glottis and the effective value of the Bernoulli effect in the glottis; (4) the effective value of the subglottic pressure; (5) the damping of the vocal folds. When these factors are known, the frequency can be calculated, using modified formulae for damped strings. The factors (3) and (4) need some further discussion.

When the other factors remain constant, the effect of the subglottic pressure on the pitch can easily be seen. For example, in Figure 9 a sudden push was exerted upon the stomach during a sustained vowel in the chest voice (the magnitude of the effect depends upon the register of voice). The pitch is seen to rise; it is very difficult to compensate for this effect of a sudden increase of the subglottic pressure. The effect can also be demonstrated with various models of the larynx. Johannes Müller had already observed it and he proposed in his 'Kompensationsgesetz der Kräfte' that a compensation is needed when the pitch is to remain constant at varying subglottic pressure. Particularly in untrained subjects the pitch is heard to rise when they are asked to make a swelling tone at constant pitch. The effect was explained (*12*) by taking into account the distension of the laryngeal cage by the increased subglottic pressure and the increased Bernoulli effect in the glottis.

The Bernoulli effect has already been mentioned by Tondorff (*50*). It is illustrated by making two papers collide when air is blown through them, and it can be demonstrated with models of the larynx. The air-resistance of the larynx and the Bernoulli effect in the glottis were recently measured by van den Berg, Zantema and Doornenbal, Jr. (*14*) by

experiments on a cast of a normal fresh human larynx, for static volume velocities up to 2 liter/second and/or subglottic pressures up to 64 cm water. Some of the results were as follows: the bulk of the air escapes under conditions of turbulent flow, whatever the form of the larynx may be, and the flow is then proportional to the area of the glottis and to the square root of the subglottic pressure. The differential resistance (that is, the resistance for variations superimposed upon the mean flow) is then of the same order of magnitude as the input resistance of the vocal cavities at the frequencies of the formants. The pressure at the outlet of the glottis is negative on account of the sucking Bernoulli effect and amounts to about half the value calculated on the assumption that energy losses are negligible; the mean pressure in the glottis effectively contributes to the stiffness of the vocal folds. The closure of the glottis is thus brought about by three basic factors: (1) the decrease of the subglottic pressure due to the escape of subglottic air, (2) the tension of the vocal folds, which causes them to become shorter, thus closing the glottis, and (3) the sucking effect of the escaping air, which is largest at the places with largest velocity of the air.

At constant mean subglottic pressure the mean flow through the larynx depends upon the mean air-resistances of larynx and vocal cavities. Normally, the resistance of the vocal cavities is much smaller than the resistance of the larynx. Under certain circumstances, however, the resistance of the vocal cavities is large enough to be significant; for example (*15*), the ventricular bands may be brought together to such an extent that they too are thrown into vibration. A parasitic vibration can also be produced at a higher level, near the tip of the epiglottis, when the epiglottis is bent towards the posterior wall of the pharynx. A large resistance can also be introduced by partially closing the nose during humming. In all these cases the fundamental frequency is lowered when the large resistance of the vocal cavities is suddenly added without compensation in the larynx. The same is seen when a vowel is alternately produced normally and throatily. These phenomena cannot be plausibly explained on the basis of the neuro-chronaxic theory, whereas they are easily explained on the basis of the myoelastic-aerodynamic theory, taking into account the increase of the supraglottal pressure and the decrease of the Bernoulli effect.

The spectrum delivered by the glottis-generator is also affected by the above procedures. A weak and relatively broad puff, as in the falsetto voice and the low intensity chest voice, will produce few higher harmonics of the fundamental tone; a strong and relatively short puff, as in the high intensity chest voice, produces a great number of high harmonics. In this respect it should be pointed out that the results of experiments on the air-resistance (*14*) of the larynx indicate that the puffs are sharper, being limited to a shorter time, than an analysis of the glottal area from high-speed motion pictures would reveal. The bulk of the air tends to escape during the first phase of large areas. Also, as the subglottic pressure drops on account of the escape of subglottic air (See Figure 3), the puff may escape in the form of a 'droplet' of air near the beginning of the opening phase. Recent experiments by Beckmann (*1*) on the primary glottal tone are in agreement with

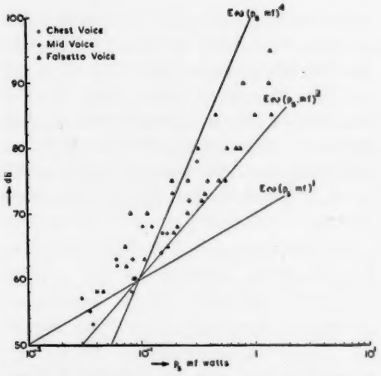

FIGURE 10. Efficiency of voice production, vowel [a]. Sound pressure at 25 cm directly in front of the mouth *versus* mean subglottic energy, mean subglottic pressure (p_s) times mean flow (mf).

these considerations. Usually, the opening of the glottis proceeds more rapidly than the closing. In a high-speed film by Vallancien, however, the closing appears more rapid than the opening and a strong glottal tone apparently originates at the closing. Presumably these differences are related to the resonance of the trachea and lungs (See Figure 2). With the chest voice the opening quotient, that is, the ratio of the duration of the opening phase to the duration of the cycle, is an important figure in this respect. With a large opening quotient, few higher harmonics are to be expected; with a small opening quotient, the number of harmonics can be very large (also with œsophageal voice). Recent experiments by Timcke (*49*) indicated that the opening quotient is rather constant at various pitches with comparable intensity of the voice. Therefore, as pointed out before (*9*, *11*), it is economical for the subject to choose his mean pitch in the high region of the pitches which he can easily produce. The efficiency of voice production has been studied (*16*) for the vowel [a] (See Figure 10). It appeared that the totally emitted sound energy, for the whole region of the voice, increased with about the fourth power of the mean subglottic pressure or mean flow and with the square of the subglottic energy. For practical purposes the spectrum of the glottis-generator in the chest voice can be taken to decrease with a slope of 6 db per octave for low frequencies and 12 db per octave for high frequencies.

So far, the properties of the vocal organs have been discussed separately; now it is necessary to investigate the effects of the coupling of these vibratory organs into a unit. Which one will determine the frequency of the fundamental tone? As a general rule, the organ with the smallest damping rules the others. Fortunately, the larynx is the system with the smallest damping. Thus man can speak or sing at the pitch he desires, at least within certain limits, without complications by the coupling of the other organs to the larynx. However, minor difficulties might be expected when the impedance of the trachea and lungs or vocal cavities approaches the impedance of the generator, that is, when the pitch comes near a resonant frequency of either trachea and lungs or vocal cavities, as the impedances are then of the same order of magnitude.

The effects of the coupling of the vocal cavities to the larynx have been described by Weiss (*55*, *56*) after his experiments with elongation of the vocal cavities by tubes taken into the mouth. The effects are then just the same as those predicted by Wien's (*57*) theory on coupling: the pitch is lowered when the formant is some-

what higher than the frequency of the vocal fold vibration and the pitch is heightened when the formant is somewhat lower. Weiss suggested that the difficulties of the register transitions were due to this type of coupling, but Trendelenburg (*51, 52, 53*) has shown that this does not apply, as the difficulties are independent of the vowel in question. Van den Berg (*17*) proved that these pitch effects arise when (1) a node of the pressures is present at the input of the vocal cavities (that is, for a formant of the vowel), and (2) when the supraglottal pressure attains a value which is comparable to the value of the subglottal pressure (that is, when the fundamental frequency approaches the formant). The supraglottal pressure is already much smaller when the second harmonic (first overtone) approaches the formant, as the volume velocity of the second harmonic is already much smaller than the volume velocity of the fundamental tone. The difficulties are then limited to rapid transitions in the strength of the second harmonic in the signal emitted from the mouth. Even with the fundamental tone approaching the formant, the difficulties are normally small, however, as the larynx has a simple means of decreasing excessive supraglottal pressures, namely, by increasing the damping of the formant by means of increased energy losses toward the trachea and lungs. This increase is brought about by a slight change of the vibratory pattern, the mean area of the glottis during the cycle becoming slightly larger.

A second type of coupling has been suggested (*17*) to explain the phase differences observed under some pathological circumstances between the vibrations of the left and right vocal folds. Phase differences and amplitude differences can arise when the tension, mass and/or the damping of the folds is unequal; the fold with the smallest damping then rules the vibration of the fold with the largest damping. Examples of this type of coupling can be seen in high-speed motion pictures by Moore and von Leden.

A third type of coupling has to be added on account of the typical vibratory behavior of the impedance of the trachea and lungs (See Figure 3). Slight difficulties are to be expected when the fundamental tone approaches the lowest resonant frequency of the trachea and connected bronchi, for example, in the neighborhood of 300 cps. These difficulties are small, however, as the larynx can slightly change its vibratory pattern to increase the damping of the subglottal system. This change might be associated with the register transition. Experiments with a tracheotomized subject are needed to study the coupling and to investigate the above possibility.

Summary

This paper reviews the evolution from the theories of Galen, Dodart and Ferrein to the present-day theories of voice production. Currently under consideration are the myoelastic-aerodynamic theory, postulating that the vocal folds are actuated by the stream of air delivered by the lungs and trachea, and the neurochronaxic theory, postulating that each vibration is induced by a nerve impulse. Anatomical, histological, physiological and physical data should provide a basis for deciding between the theories. The properties of the trachea-lungs system, the larynx and the vocal cavities, and the coupling

of these organs to form a unit are discussed. Attention is drawn to two new types of coupling, that of the vocal folds to each other and that of the trachea-lungs system to the larynx. It is shown that the myoelastic-aerodynamic theory provides a straightforward explanation of all known phenomena of voice production, whereas there is no experimental evidence for the neuro-chronaxic theory and it is unable to explain a large number of phenomena.

References

1. Beckmann, G., Elektroakustische Untersuchungen über den primären Kehlkopfton. *Arch. Ohr.-, Nas.-, u. Kehlkheilk.*, 169, 1956, 196-201.
2. Behringer, S., Die Anordnung der Muskulatur in der menschlichen Stimmlippe und im Gebiet des Conus elasticus. *Z. Anat.*, 118, 1955, 324-342.
3. Berg, J. van den, Transmission of the vocal cavities. *J. acoust. Soc. Amer.*, 27, 1955, 161-168.
4. Berg, J. van den, Calculations on a model of the vocal tract for vowel /i/ (meat) and on the larynx. *J. acoust. Soc. Amer.*, 27, 1955, 332-338.
5. Berg, J. van den, to be published.
6. Berg, J. van den, to be published.
7. Berg, J. van den, Physica van de stemvorming, met toepassingen. Doctoral thesis, Groningen, 1953.
8. Berg, J. van den, Physiology and physics of voice production. *Acta physiol. pharm. néerl.*, 5, 1956, 40-55.
9. Berg, J. van den, Physiologie et physique de la vibration des cordes vocales. *Larynx et Phonation*, Paris: Presses Universitaires de France, 1957.
10. Berg, J. van den, and Moll, J., Zur Anatomie des menschlichen Musculus vocalis. *Z. Anat.*, 118, 1955, 465-470.
11. Berg, J. van den, Sur les théories myoélastique et neuro-chronaxique de la phonation. *Rev. Laryng., Bordeaux*, 74, 1954, 494-512.
12. Berg, J. van den, Subglottic pressures and vibrations of the vocal folds. *Folia phoniatr.*, 9, 1957, 65-71.
13. Berg, J. van den, and Spoor, A., Microphonic effect of the larynx. *Nature*, 179, 1957, 625-626.
14. Berg, J. van den, Zantema, J. T., and Doornenbal Jr., P., On the air resistance and the Bernoulli effect of the human larynx. *J. acoust. Soc. Amer.*, 29, 1957, 626-631.
15. Berg, J. van den, On the role of the laryngeal ventricle in voice production. *Folia phoniatr.*, 7, 1955, 57-69.
16. Berg, J. van den, Direct and indirect determination of the mean subglottic pressure. *Folia phoniatr.*, 8, 1956, 1-24.
17. Berg, J. van den, Über die Koppelung bei der Stimmbildung. *Z. Phonet.*, 8, 1954, 281-293.
18. Dodart, D., Sur les causes de la voix de l'homme, et de ses différens tons. *Hist. Acad. roy. Sc.*, 1700, 244-274-293.
19. Dodart, D., Sur la voix et sur les tons. *Hist. Acad. roy. Sc.*, 1706, 136-148, 388-410.
20. Dodart, D., Sur la voix et les tons. *Hist. Acad. roy. Sc.*, 1707, 66-81.
21. Dunn, H. K., The calculation of vowel resonances, and an electrical vocal tract. *J. acoust. Soc. Amer.*, 22, 1950, 740-753.
22. Faaborg-Andersen, K., Electromyographic investigations of intrinsic laryngeal muscles in humans. Thesis, Copenhagen, 1957.
23. Ferrein, A., De la formation de la voix de l'homme. *Hist. Acad. roy. Sc.*, 1741, 409-432.
24. Fessard, A., and Vallancien, B., Résultats de la stimulation du nerf récurrent, avec électromyographie des cordes vocales chez le chien. *J. Physiol. Path. gén.*, 48, 1956, 538-541.
25. Fessard, A., and Vallancien, B., Observations et enregistrements électromyographiques au cours de l'activité phonatoire provoquée ou spontanée chez le chien. In *Larynx et Phonation*, Paris: Presses Universitaires de France, 1957.
26. Fessard, A., and Vallancien, B., Données électrophysiologiques sur le fonctionnement de l'appareil phonatoire du chien. *Folia phoniatr.*, 9, 1957, 152-163.
27. Goerttler, K., Die Anordnung, Histologie und Histogenese der quergestrieften Muskulatur im menschlichen Stimmband. *Z. Anat.*, 115, 1950, 352-401.
28. Guillemin, A., Sur la génération de la voix et du timbre. Paris: Société d'études scientifiques, 1897.
29. Helmholtz, H. von, Die Lehre der Tonempfindungen als Physiologische grundlage für die Theorie der Musik. Braunschweig: F. Vieweg u. Sohn, 1863.

30. Heymann, O., Das Phänomen des aktiven Schwingungsimpulses der Stimmlippen sowie die neuro-chronaxischen und cerebralen Beobachtungen Hussons in ihrer Bedeutung für die Phonation gesunder und kranker Stimmen. *Arch. Ohr.-, Nas.-, u. Kehlkheilk.*, 164, 1953, 190-196.
31. Husson, R., Étude des phénomènes physiologiques et acoustiques fondamentaux de la voix chantée. Thesis, Paris, 1950.
32. Husson, R., Sur la physiologie vocale. *Ann. Oto-laryng.*, 69, 1953, 124-137.
33. Husson, R., Physiologie de la phonation et phoniatrie. *Rev. Quest. sci.*, 1954, 75-89.
34. Husson, R., Physiologie de la phonation, in *Encycl. Medico Chirurg.*, Paris: 1955.
35. Laget, P., Réproduction expérimentale de la vibration des cordes vocales, en l'absence de tout courant d'air, par stimulation électrique d'un recurrent du chien, avec observation stroboscopique de la résponse laryngée. *Rev. Laryng., Bordeaux*, 74, 1953, 132-142.
36. Maatz, R., Die Atemstütze im Kunstgesang. *Arch. Sprach-u. Stimmheilk.*, 1, 1937, 110-127.
37. Mayet, A., Bau und Funktion des M. vocalis und seine Beziehungen zu Lig. Vocale und Conus elasticus. *Acta Anat.*, 24, 1955, 15-26.
38. Mayet, A., Zur funktionellen Anatomie der menschlichen Stimmlippe. *Z. Anat.*, 119, 1955, 87-111.
39. Moulonguet, A., Laget, P., and Husson, R., Démonstration, chez l'homme, de l'existence dans le nerf récurrent de potentiels d'action moteurs synchrones avec les vibrations des cordes vocales. *Bull. Acad. Nat. Méd.*, 137, 1953, 475-482.
40. Piquet, J., Decroix, G., Étude expérimentale peropératoire du rôle de la pression sous glottique sur la vibration des cordes vocales. *Comptes rend. Acad. Sc.*, 242, 1956, 1223-1225.
41. Portmann, G., Humbert, R., Robin, J. L., Laget, P., and Husson, R., Étude électromyographique des cordes vocales chez l'homme. *Comptes rend. Soc. Biol.*, 149, 1955, 296-300.
42. Portmann, G., Robin, J. L., Laget, P., and Husson, R., La myographie des cordes vocales. *Acta Otolaryng., Stockh.*, 46, 1956, 250-263.
43. Savart, F. Mémoire sur la voix humaine. *Ann. Chim.* (*Phys.*) par Gay Lussac et Arago, 30, 1825, 64-87.
44. Smith, S., Chest register versus head register in the membrane cushion model of the vocal cords. *Folia phoniatr.*, 9, 1957, 32-36.
45. Smith, S., Remarks on the physiology of the vibrations of the vocal cords. *Folia phoniatr.*, 6, 1954, 166-178.
46. Sonninen, A., The role of the external laryngeal muscles in length adjustment of the vocal cords in singing. Thesis, Helsinki, 1956.
47. Spoor, A., and Dishoeck, A. H. E. van, to be published.
48. Stevens, K. N., Kasowski, S., and Fant, C. G. M., An electrical analog of the vocal tract. *J. acoust. Soc. Amer.*, 25, 1953, 734-742.
49. Timcke, R., Die Synchron-Stroboskopie von menschlichen Stimmlippen bzw ähnlichen Schallquellen und Messung der Öffnungszeit. *Z. Laryng. Rhinol.*, 35, 1956, 331-335.
50. Tonndorff, W., Die Mechanik bei der Stimmlippenschwingung und beim Schnarchen. *Z. Hals-, Nas.-, u. Ohrenheilk.*, 12, 1925, 241-245.
51. Trendelenburg, W., Untersuchungen zur Kenntnis der Registerbruchstellen beim Gesang. Sitz, Ber. *Preuss. Akad. Wiss.*, 1938, 5-22, 188-226.
52. Trendelenburg, W., Zur Frage der Koppelungsbeziehungen zwischen Ansatzraum und Stimmlippenmechanismus. *Proc. 3 Int. Congr. Phonet. Sc.* (Ghent), 1938, 188-190.
53. Trendelenburg, W., Über die Frage der Koppelung zwischen Ansatzrohr und Windrohr beim menschlichen Stimmorgan Abh. *Preuss. Akad. Wiss.*, 1940, nr. 9.
54. Vosteen, K. H., Récherches histologiques en lumière polarisée sur la musculature des cordes vocales. *Rev. Laryng. Bordeaux*, 78, 1957, 608-616.
55. Weiss, D., Ein Rosonanzphänomen der Singstimme. *Mschr. Ohrenheilk.*, 66 1932, 964-967.
56. Weiss, D., Zur Frage der Registerbruchstellen. *Z. Hals-, Nas.-, u. Ohrenheilk.*, 30, 1932, 353-358.
57. Wien, M., Über die Rückwirkung eines resonirenden Systems. *Ann. Phys. Chem.* (N.F.), 61, 1897, 151-189.
58. Wustrow, F., Bau und Function des menschlichen Musculus vocalis. *Z. Anat.*, 116, 1952, 506-522.

Infant Speech: Consistency With Age

Harris Winitz

Orvis C. Irwin

Within the past ten years research conducted at the Iowa Child Welfare Research Station has led to a more precise understanding of the early speech development of infants. Irwin and his associates studied longitudinally over a period of two and one-half years the speech sound utterances of 95 white infants (*4, 5, 6, 7, 8, 9, 10, 11, 12*). In the present investigation correlational procedures were applied to certain selected measures of speech sound development to determine the extent to which these same infants maintained their positional standing in the group with increase in age.

The infants were from Iowa City homes and were considered to be physically normal, having a median birth weight of seven and one-half pounds. With few exceptions the children were from monolingual homes. The socio-economic status of the group was described by Irwin (7) as follows: 'These infants were from middle class homes, the parents being professional, business, clerical, and some laboring people.' The infants were divided approximately equally with respect to sex and family constellation. With regard to family constellation 'only children' were defined as having no older siblings, while 'infants with siblings' had an older sibling; the age interval between the child and his sibling was not considered.

As the study progressed, some of the families moved away so that fewer subjects were available at the older age levels. Irwin calls 'most continuous data' those data obtained from the 35 subjects who were most continuously followed throughout the two-and-one-half-year period of infancy. He refers to the data obtained from all 95 subjects as 'all data.'

The data were collected in the following manner: The spontaneous speech sounds of the infants (no attempt was made by the recorder to stimulate the child verbally) were transcribed on paper using the abbreviated International Phonetic Alphabet described by Fairbanks (*3*). Two vowel pairs, [ə]—[ʌ] and [æ]—[a], were considered as two vowels, and three consonants were added: [x], [ʔ] and [ç]. Lip smacking sounds and 'clicking' sounds were omitted in the transcriptions.

The sampling unit was 30 breaths; that is, the sounds which were made by an infant on 30 exhalations, not necessarily consecutive, were recorded. Transcriptions were limited to non-crying sounds. A transcription of

Harris Winitz (M.A., University of Iowa, 1956) is a graduate assistant in the Department of Speech Pathology and Audiology, University of Iowa, and Orvis C. Irwin (Ph.D., Ohio State University, 1929) is Professor of Psychology, Iowa Child Welfare Research Station.

30 breaths was called a 'record'; each record took approximately 30 minutes to collect. This sampling procedure was found to have high observer reliability (*4*). All the data were collected during the afternoon after the child had his noon meal. No attempt was made to control the infant's position during the making of a record, although almost all were sitting or held in an upright position. Usually one parent, the mother, was present. The data were grouped and reported in 15 two-month intervals or age levels, beginning with months one and two (age level one) and ending with months 29 and 30 (age level 15).

Problem

Bayley (*1*) and Shirley (*13*) studied the relationship between the acquisition of early and late motor skills in an attempt to discover whether children who were advanced in motor skills at an early age maintained their rank at later age levels. The correlation coefficients they report suggest a high relationship between motor events close in time, such as walking with help and walking alone, but indicate only a slight relationship between the attainment of early and late motor skills, such as reaching for and grasping an object and walking.

The present investigation was undertaken to examine a similar relationship with respect to speech sound development. The purpose was to determine the extent to which the infant maintains his rank in the group (with respect to speech sound development) from early to late age levels.

Procedure

The original data collected by Irwin and his associates were used for this study. Two indices of speech

TABLE 1. Phoneme type: Pearson product-moment correlation coefficients for phoneme type computed for each pair of odd-numbered age levels (two month groupings) from age level one (months one and two) to age level 15 (months 29 and 30). The number of subjects for each correlation is indicated in parentheses.

Age Levels	*Age Levels*						
	3	*5*	*7*	*9*	*11*	*13*	*15*
1	.04 (38)	.13 (35)	−.07 (38)	−.06 (28)	.13 (23)	−.10 (20)	.03 (12)
3		.21 (47)	.05 (49)	.09 (43)	.06 (31)	−.04 (27)	−.18 (17)
5			.26 (48)	.43* (44)	.29 (32)	.08 (29)	.17 (18)
7				.08 (44)	.12 (34)	.37† (31)	−.02 (18)
9					.38† (35)	.27 (30)	−.01 (18)
11						.27 (29)	.32 (18)
13							.16 (18)

*Significant at the one per cent level.
†Significant at the five per cent level.

TABLE 2. Phoneme frequency: Pearson product-moment correlation coefficients for phoneme frequency computed for each pair of odd-numbered age levels (two month groupings) from age level one (months one and two) to age level 15 (months 29 and 30). The number of subjects for each correlation is indicated in parentheses.

Age Levels	*Age Levels*						
	3	*5*	*7*	*9*	*11*	*13*	*15*
1	.28	.40†	.25	.15	.35	.26	.38
	(38)	(35)	(38)	(28)	(23)	(20)	(12)
3		.53*	.04	.25	.11	.12	−.10
		(47)	(49)	(43)	(31)	(27)	(17)
5			.46*	.34†	.16	.50*	.43
			(48)	(44)	(32)	(29)	(18)
7				.33†	.26	.29	.04
				(44)	(34)	(31)	(18)
9					.44*	.63*	.48†
					(35)	(30)	(18)
11						.22	.43
						(29)	(18)
13							.47†
							(18)

*Significant at the one per cent level.
†Significant at the five per cent level.

sound development were selected: phoneme type and phoneme frequency. Phoneme type refers to the number of different sounds in the record of a given child; phoneme frequency refers to the number of phonemes, regardless of type, in the record of a given child. Pearson product-moment correlations for phoneme type and phoneme frequency were computed for each pair of odd-numbered age levels from one to 15. Since the data were not entirely continuous, the number of subjects available for the inter-age level correlations varied. In this study sexes were grouped together.

Results

The obtained correlation coefficients for phoneme type are shown in Table 1. The relationships, for the most part positive, are quite low. Only three coefficients (between age levels five and nine, seven and 13, and nine and 11) are statistically significant at or beyond the five per cent level. It is important to note that correlations computed between later age levels would tend to be depressed because of the arbitrary ceiling (namely, the number of possible phoneme types that can be learned) placed on the English-speaking child.

Ten correlation coefficients were found to be significantly different from zero for phoneme frequency (Table 2). The highest coefficient obtained was .63 (between age levels nine and 13) and the lowest .33 (between age levels seven and nine). Only three of these coefficients exceed .50, indicating, in general, a very slight positive relationship between age levels. It is apparent from an examination of the two tables that

phoneme frequency yields more significant results than phoneme type. It should be pointed out that in contrast to the limited ceiling on phoneme type, the ceiling on phoneme frequency is practically unlimited.

A further examination of Tables 1 and 2 indicates that statistically significant correlations are found, in the main, between age levels close in time. Perhaps if coefficients were computed between successive age levels, such as three and four, additional statistically significant and higher correlations would be found.

Discussion

It appears from the results of this study that the infants under investigation, in general, did not maintain their relative positions in the group from odd-numbered age level to odd-numbered age level with respect to phoneme type and phoneme frequency measures of speech sound development. The highest relationships were found to occur between age levels close in time, a result which has been similarly found in studies of motor development (*1*, *13*).

A word of caution regarding the interpretation and use of the Irwin data is in order. The lack of stability of these measures from one age level to another would appear to limit their usefulness in at least two ways. First, doubt is cast on the feasibility of using these measures for *predictive* purposes of the speech sound status of an infant during the first two and one-half years (the terminal point of the data). This does not mean that studies of infant speech vocalizations, analyzed in other ways, might not serve useful *diagnostic* purposes. The use of consonant and vowel profiles and percentile grids previously suggested by Irwin (*6*, *8*, *9*) may yet prove to be useful clinical tools. Second, speech sound measures for the ages studied would probably have little use as predictors of performance on intelligence tests at later ages. Catalano and McCarthy (*2*) have suggested speech sound measures as a possible predictor of later intelligence. The results of the present study, however, suggest that such measures are probably too unstable from one age level to another to serve, in general, as reliable predictors of later performance of any kind. Extreme values, of course, either high or low, might be an exception.

Summary

The consistency of infant speech vocalizations with respect to age was investigated. From data on file at the Iowa Child Welfare Research Station, inter-age level Pearson product-moment correlation coefficients for phoneme type and phoneme frequency were computed for each pair of odd-numbered age levels (two month groupings) from one (months one and two) to 15 (months 29 and 30). Of the 56 correlation coefficients obtained, only 13 were statistically significant and in only three instances did the coefficients exceed .50. It appears, then, that the infants under investigation in this study did not consistently maintain their positional standing in the group from odd-numbered age level to odd-numbered age level with respect to the selected measures of speech sound development. The correlation coefficients are not of sufficient magnitude or significance to warrant their use for predictive purposes.

References

1. Bayley, N., The development of motor abilities during the first three years. *Society for Research in Child Development, Monogr.*, 1, 1935, No. 1.
2. Catalano, F. L., and McCarthy, D., Infant speech as a possible predictor of later intelligence. *J. Psychol.*, 38, 1954, 203-209.
3. Fairbanks, G., *Voice and Articulation Drillbook*. New York: Harpers, 1940.
4. Irwin, O. C., Reliability of infant speech sound data. *J. Speech Dis.*, 10, 1945, 227-235.
5. Irwin, O. C., and Chen, H. P., Infant speech sounds and intelligence. *J. Speech Dis.*, 10, 1945, 293-296.
6. Irwin, O. C., Infant speech: variability and the problem of diagnosis. *J. Speech Dis.*, 12, 1947, 287-289.
7. Irwin, O. C., and Chen, H. P., Development of speech during infancy: curve of phonemic types. *J. exp. Psychol.*, 36, 1946, 431-436.
8. Irwin, O. C., Infant speech: consonantal sounds according to place of articulation. *J. Speech Dis.*, 12, 1947, 397-401.
9. Irwin, O. C., Infant speech: development of vowel sounds. *JSHD*, 13, 1948, 31-34.
10. Irwin, O. C., Infant speech: the effect of family occupational status and of age on use of sound types. *JSHD*, 13, 1948, 224-226.
11. Irwin, O. C., Infant speech: the effect of family occupational status and of age on sound frequency. *JSHD*, 13, 1948, 320-323.
12. Irwin, O. C., Infant speech: speech sound development of sibling and only infants. *J. exp. Psychol.*, 38, 1948, 600-602.
13. Shirley, M. M., *The First Two Years: A Study of 25 Babies. Vol. 1., Postural and Locomotor Development*. Minneapolis: Univ. of Minnesota Press, 1931.

Syllabic And Phonetic Structure Of Infants' Early Words

Harris Winitz

Orvis C. Irwin

The purpose of this article is to report an analysis of three aspects of infant speech, namely, the syllabic structure, the phonetic structure and the vowel and consonant composition of early words. The data were selected from files of the Iowa Child Welfare Research Station. A complete account of the subjects and of the methodology employed in the original collection and analysis of the data has been described by Winitz and Irwin (*13*). With regard to this present investigation an additional fact should be mentioned: phonetic utterances interpreted as words were recorded both in the International Phonetic Alphabet and in the English alphabet.

Procedure

Definition of a Word. In this study an utterance was counted as a *word* if it could be categorized as: (1) a word approximation, (2) a standard word or (3) a self-language word.

Harris Winitz (M.A., University of Iowa, 1956) is a graduate assistant in the Department of Speech Pathology and Audiology, University of Iowa, and Orvis C. Irwin (Ph.D., Ohio State University, 1929) is Professor of Psychology, Iowa Child Welfare Research Station.

Word Approximation. In a recent study McCurry and Irwin (*11*) called attention to the fact that the infant's early words may not be of the same phonetic structure as those of adults but may be approximations of words in which some of the phonemes are the same. They have defined a *word approximation* as 'a phonetic pattern which is interpreted by the observers at the time of the transcription as an attempt by the infant to pronounce a standard word. The word approximation is further delimited as a phonetic pattern in which one or more of the phonetic elements of the standard word, either vowel or consonantal elements, are present. This means that some elements of the standard word are omitted, and other elements are substituted or added.'

Standard Word. McCurry and Irwin (*11*) have defined a *standard word* in terms of its phonetic listing in Kenyon and Knott's *Pronouncing Dictionary of American English* (*5*). The present investigators felt that several exceptions to the definition of a standard word were justified because many words used in the home by parents and learned by children are not of the exact phonetic structure listed by Kenyon and Knott. The

following gives examples of such exceptions:

> The following form of 'daddy' was accepted: [dædi]. The following forms of 'mother' were accepted: [mɑmi], [mɑmɑ], [mɑmʌ], [mɑ], [mɑm], and [mɑmə]. The following forms of 'baby' were accepted. [beibi] and [bebi]. Words such as 'doggy,' 'choo-choo,' 'bow wow,' 'tick-tock,' 'moo moo,' etc., were considered as standard words when they were of the correct phonetic structure. If these words were not of the correct phonetic structure, they were considered word approximations.

Self-Language Word. Another category has been added to include some of the infant's first words which do not meet the criteria of standard word or word approximation. This category, *self-language word*, may be defined as a phonetic pattern which is interpreted by the observers at the time of transcription as an attempt by the infant to pronounce a standard word, although not one of the phonetic elements of the standard word is present. An instance of this is [bʌ] for 'dog.'

Selection of Data. The data, as described by Winitz and Irwin (*13*), were grouped and reported in 15 two-month intervals or age levels, beginning with age level one (months one and two) and ending with age level 15 (months 29 and 30). Age levels seven (months 13 and 14), eight (months 15 and 16) and nine (months 17 and 18) were selected for analysis in this study because they represent, according to previous investigations (*10*), the period of infancy during which infants are reported to utter their first words. The records selected for this study were those for children who, according to the available records, uttered at least one word at one of the above age levels.

TABLE 1. Number of subjects, records and words at age levels seven (months 13 and 14), eight (months 15 and 16) and nine (months 17 and 18).

Age Level	*Subject*	*Records*	*Words*
7	23	38	137
8	35	62	224
9	35	58	268

Table 1 indicates the distribution of subjects, records and words at each age level. Table 2 shows the number and percentage of standard words, word approximations, and self-language words that were subject to analysis at each age level.

Method of Analysis. In line with the specific purposes of this paper the

TABLE 2. Number and percentage of standard words, word approximations, and self-language words in the speech of infants at age levels seven (months 13 and 14), eight (months 15 and 16) and nine (months 17 and 18).

Age Level	*Standard Words*		*Word Approximations*		*Self-Language*	
	N	*%*	*N*	*%*	*N*	*%*
7	22	16.06	114	83.21	11	.73
8	53	23.66	170	75.89	1	.45
9	103	38.43	151	56.35	14	5.22

TABLE 3. Percentages of word types in the speech of infants at age levels seven (months 13 and 14), eight (months 15 and 16) and nine (months 17 and 18). Words were classified as (1) monosyllabic words with two phonemes; (2) monosyllabic words with three phonemes; (3) words consisting of a repeated syllable; (4) words consisting of a repeated syllable with an added phoneme on the final syllable; (5) words consisting of two syllables in which one sound is common and one sound varies; and (6) words consisting of two syllables in which one sound is common and one sound varies and a phoneme is added to the second syllable.

Age Level	*Word Classification*						*Total*
	1	*2*	*3*	*4*	*5*	*6*	
7	18.25	5.84	17.52	2.92	31.39	0.00	75.92
8	44.20	20.09	5.36	0.00	15.63	1.34	86.62
9	34.33	15.67	9.33	.37	28.36	.75	88.81

data were analyzed for each of the following categories: (1) syllabic structure, (2) phonetic structure and (3) vowel and consonant composition.

Syllabic Structure. In order to analyze the syllabic structure of the infant's early words, a modification of the categories employed by Chen (2), in a study of the repetitious patterns of the phonemes of infants, was used and the following categories were employed: (1) monosyllabic words with two phonemes, such as [bi]; (2) monosyllabic words with three phonemes, such as [bim]; (3) words consisting of a repeated syllable, such as [bibi]; (4) words consisting of a repeated syllable with an added phoneme on the final syllable, such as [bibim]; (5) words consisting of two syllables in which one sound is common and one sound varies, such as [bibɪ]; (6) words consisting of two syllables in which one sound is common and one sound varies and a phoneme is added to the second syllable, such as [bibɪm]. In the above analysis diphthongs, blends, and affricates were considered single units or 'phonemes' since these combinations of sounds tend to be heard as elemental units.

TABLE 4. Vowel phoneme percentages in the words of infants at age levels seven (months 13 and 14), eight (months 15 and 16) and nine (months 17 and 18).

Vowel	*Age Level*		
	7	*8*	*9*
[i]	5.49	13.57	13.79
[ɪ]	11.76	10.80	7.14
[e]	5.49	8.03	5.81
[ɛ]	15.69	4.43	4.49
[æ]	7.06	7.48	5.81
[ʌ]	18.04	10.52	8.80
[ə]	1.18	.83	.17
[ɑ]	19.22	18.28	31.89
[ɔ]	.78	1.66	1.10
[o]	2.75	6.88	6.31
[ʊ]	8.63	7.20	9.30
[u]	3.92	10.53	5.15

Phonetic Structure. To investigate the phonetic structure of the infant's early words, the elemental sounds or phonemes were tallied and presented in terms of percentages. Additional analyses included several vowel (front, middle and back) and consonantal (by place and manner of articulation) categories.

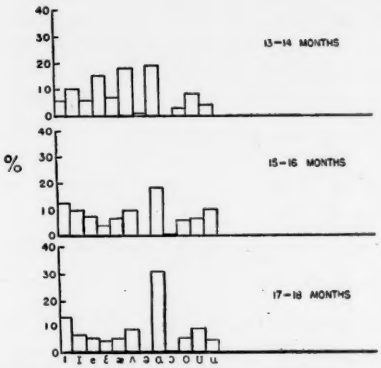

FIGURE 1. Profiles of vowel phoneme percentages in the words of infants at age levels seven (months 13 and 14), eight (months 15 and 16) and nine (months 17 and 18).

Consonant and Vowel Composition. To investigate the consonant-vowel composition of infants' early words, the words were analyzed to determine whether they were constituted only of vowels, only of consonants, or of both vowels and consonants.

Results

Syllabic Structure. The results of the analysis of the syllabic structure of the words are shown in Table 3. An examination of the last column in this table indicates that most of the infant's early words are monosyllables or dissyllables as they were classified for the purposes of this analysis. These types of words seem to vary in percentage of use, however, with age. The obtained results were 76% at age level seven, 87% at age level eight and 89% at age level nine.

Phonetic Structure. Table 4 indicates what percentage each vowel comprises of the total number of vowels used in words for each age

TABLE 5. Percentage of vowel categories, classified according to Fairbanks (3), used in the words of infants at age levels seven (months 13 and 14), eight (months 15 and 16) and nine (months 17 and 18).

Age Level	*Front*	*Middle*	*Back*
7	45.49	19.21	35.30
8	44.31	11.63	44.32
9	36.87	9.30	53.82

TABLE 6. Consonant phoneme percentages in the words of infants at age levels seven (months 13 and 14), eight (months 15 and 16) and nine (months 17and 18).

Consonant	*Age Level* 7	8	9
[p]	.42	2.44	1.15
[b]	33.90	27.74	37.61
[m]	17.80	7.93	5.28
[w]	12.29	4.88	4.13
[ʍ]	.00	.00	.00
[f]	.00	.30	3.21
[v]	.00	.00	.00
[θ]	.00	.00	.00
[ð]	.42	.00	.46
[t]	2.12	7.62	4.13
[d]	23.31	19.82	19.27
[n]	.85	4.57	7.11
[s]	.85	2.44	2.06
[z]	.00	1.52	.92
[ʃ]	.00	.61	1.83
[ʒ]	.00	.00	.00
[l]	1.27	.30	1.37
[r]	.00	2.44	.97
[j]	.00	1.83	.00
[ç]	.00	.00	.00
[ŋ]	.00	.00	.23
[k]	2.54	8.23	4.36
[g]	2.12	3.36	4.36
[x]	.00	.00	.00
[h]	2.12	3.96	1.61
[ʔ]	.00	.00	.00

level. Thus the vowel [i], for example, comprises 5.49% of the total number of vowels used during the seventh age level, 13.57% during the eighth and 13.79% at the ninth. Vowel profiles for the three age levels are shown in Figure 1. The profile for age level seven reveals that the vowels [ɪ], [ɛ], [ʌ] and [ɑ] are those most frequently used in words. By the ninth age level the vowel [ɑ] shows a great increase in percentage of occurrence in word usage. The sound [ɪ] in the ninth age level has doubled in percentage of usage in words over the seventh age level. The [ɛ] and [ʌ] vowels show some decrease in percentage of usage with increased age.

Table 5 indicates the percentages of vowel categories, classified according to Fairbanks (*3*), used in the words of infants at age levels seven through nine. This table shows that the percentage of front vowel usage in words is greater at age level seven than the percentage of back vowel usage. At age level eight the percentages are the same, but at age level nine the percentage of back vowels is considerably higher than that of the front vowels. The use of the middle vowels relative to other vowels decreases from age level seven to age level nine.

The percentages of consonantal phonemes are listed in Table 6. Consonant profiles appear in Figure 2. An examination of the table and the profiles indicates that labials are used in words more frequently than are other consonants. Among the labials the [b] sound is dominant. The postdentals as a group are next in order of usage in words. Here the consonant [d] is the outstanding phoneme. The profiles show that back consonants at the seventh age level comprise a small percentage of the total consonants used in words. The percentage of back consonants becomes slightly higher at

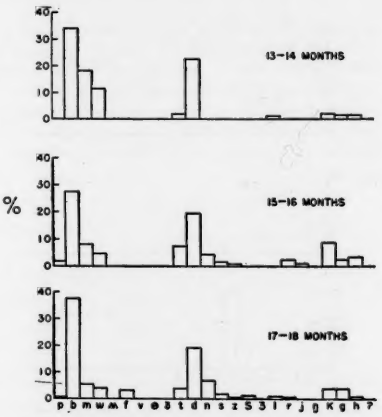

FIGURE 2. Profiles of consonant phoneme percentages in the words of infants at age levels seven (months 13 and 14), eight (months 15 and 16) and nine (months 17 and 18).

TABLE 7. Percentages of consonants, classified according to Fairbanks (*3*) by place of articulation, used in the words of infants at age levels seven (months 13 and 14), eight (months 15 and 16) and nine (months 17 and 18).

Age Level	*Labial*	*Labio-Dental*	*Lingua-Dental*	*Post-Dental*	*Velar*	*Glottal*
7	64.41	.00	.42	28.40	4.66	2.12
8	42.99	.30	.00	41.15	11.59	3.96
9	48.17	3.21	.46	37.61	8.95	1.61

TABLE 8. Percentages of consonants, classified according to Fairbanks (*3*) by manner of articulation, used in the words of infants at age levels seven (month 13 and 14), eight (months 15 and 16) and nine (months 17 and 18).

Age Level	*Nasal*	*Plosive*			*Semi-vowel*	*Fricative*			*Glide*
		Voiced	*Voiceless*	*Total*		*Voiced*	*Voiceless*	*Total*	
7	18.65	59.33	5.08	64.41	1.27	.42	2.97	3.39	12.29
8	12.50	50.92	18.29	69.21	2.74	1.52	7.31	8.83	6.71
9	12.62	61.24	9.64	70.88	2.29	1.38	8.71	10.09	4.13

age levels eight and nine. This finding indicates that when words appear in the vocalization of infants a reversal has occurred in the situation which Irwin (*4, 5*) found concerning the status of back consonants during the first months of life. During early infancy velars and glottals are relatively frequent, but with the appearance of words during the second year of infancy there is a decrement in the relative frequency of their usage.

Tables 7 and 8 indicate the percentages of consonants, classified according to Fairbanks (*3*) by place and manner of articulation, used at the three age levels. With regard to place of articulation the labial and postdental sounds constitute more than 80% of the consonantal sounds at each age level. With regard to manner of articulation the plosive consonants greatly predominate over nasal, semivowel, fricative and glide consonants at all three age levels.

Vowel and Consonant Composition. Table 9 shows that about 95% of the early words of infants contain both vowels and consonants. At age level seven only a small percentage of words composed only of consonants were recorded. At all age levels words composed only of vowels constitute about five per cent or less of the total words spoken.

Discussion

Syllabic Structure. The findings of this study appear to be in agreement with reports found in the literature. Berry and Eisenson (*1*) have suggested that the syllabic structure of the first words are either dissyllables, such as *baba*, or monosyllables, such as *ba*. Similarly, McCarthy (*10*) has stated, 'Children's babbling often consists of repetition of identical or similar syllables so that the first vocal utterances to acquire meaning are usually reduplicated monosyllables such as *mama, dada, nana, bye-bye,*

TABLE 9. Percentages of words composed only of vowels, only of consonants, or of both vowels and consonants in the speech of infants at age levels seven (months 13 and 14), eight (months 15 and 16) and nine (months 17 and 18).

Age Level	*Vowels*	*Consonants*	*Vowels and Consonants*
7	5.11	.73	94.16
8	4.45	.00	95.55
9	3.36	.00	96.64

tick-tick, choo-choo, and the like.' Lewis (*8*) has reported that 85% of early words are either monosyllabic or reduplicated monosyllables (39% of the former and 46% of the latter).

Phonetic Structure. This study has some bearing on the controversy concerning which consonants are first mastered by the infant, namely, whether they are labials or back consonants. The statement is made in the literature that labials are the earliest sounds (*6, 9, 12*). The Iowa studies (*4, 5*) have shown that in the early months of life the first consonants uttered are back consonants and that labials are relatively infrequent. The present study has shown that during the second year of life labial consonants are most frequent as elements in word structure. The inference seems reasonable, then, that some of the observations reported previously in the literature were not made on infants in the early months of life, but were made on infants who were beginning to say their first words.

Summary

The syllabic structure, phonetic structure and vowel and consonant composition of the infant's early words were studied at three age levels: the seventh (months 13 and 14), the eighth (months 15 and 16), and the ninth (months 17 and 18).

A high percentage of the words were either monosyllables or dissyllables. The vowel sounds varied in relative use at the different age levels with the exception of the vowel [ɑ], which was outstanding at each age level. The labial and post-dental sounds constituted more than 80% of the consonant sounds at each age level. A higher percentage of front and back vowels than of middle vowels were used. Approximately 95% of the words were composed of both vowels and consonants.

References

1. Berry, M. F., and Eisenson, J. *Speech Disorders; Principles and Practices of Therapy.* New York: Appleton-Century-Crofts, Inc., 1956.
2. Chen, H. P., Speech development during the first year of life, a quantitative study. Ph.D. dissertation, University of Iowa, 1946.
3. Fairbanks, G., *Voice and Articulation Drillbook.* New York: Harpers, 1940.
4. Irwin, O. C., Infant speech: consonant sounds according to place of articulation. *J. Speech Dis.,* 12, 1947, 397-401.
5. Irwin, O. C., Infant speech: consonant sounds according to manner of articulation. *J. Speech Dis.,* 12, 1947, 402-404.
6. Jesperson, J. O. H., *Language: Its Nature, Development and Origin.* New York: Holt, 1928.
7. Kenyon, J. S., and Knott, T. A., *A Pronouncing Dictionary of American English.* Springfield, Mass.: Merriam Co., 1949.
8. Lewis, M. M., *Infant Speech: a Study of the Beginnings of Language* (2nd ed.). London: Routledge and Paul, 1951.
9. Löbisch, J. E. L., *Entwickelungsgeschicte der Seele des Kindes.* Wein: C. Haas, 1851.
10. McCarthy, D., Language development in children, in *Manual of Child Psychology, Ed. by L. Carmichael* (2nd ed.). New York: John Wiley & Sons, Inc., 1954.
11. McCurry, W. H., and Irwin, O. C., A study of word approximations in the spontaneous speech of infants. *JSHD,* 18, 1953, 133-139.
12. Tracy, F., *The Psychology of Childhood* (3rd ed.). Boston: Heath, 1896.
13. Winitz, H., and Irwin, O. C., Infant speech: consistency with age. *JSHR,* 1, 1958, pp. 245-249.

Application Of Descriptive Linguistics To Child Language

Robert W. Albright

Joy Buck Albright

Modern linguistic science has two chief branches, one of descriptive and the other of historical linguistics. The study of child language, from a linguistic point of view, is considered to be a special branch of historical linguistics (*8*). In applying linguistic methods to child language, however, it is important to be aware of the differences between it and conventional language. A child's language is not used by a community of speakers, and it is changing rapidly, in contrast to the slow rate of change in conventional languages. Consequently, the collection and analysis of specimens of child language require a special adaptation of the techniques of descriptive linguistics. Some adaptation of techniques is necessary, of course, in any linguistic study, and Wells (*9*) has made the useful suggestion that 'In descriptive linguistics, discovery consists in finding the best scheme in terms of which to describe the facts . . .'

One important adaptation is to substitute what might be called the 'equivalence of recurrence' for the linguist's 'equivalence of repetition.' Not having adult informants who can conveniently repeat forms for comparison as being the 'same' or 'different,' linguists working with children often depend upon the forms recurring at different times in a sample of discourse. Thus, with an adult speaker of English, it would not be hard to contrast speech forms such as *you*, *her*, and *him* in *I see you*, *I see her* and *I see him*. But if a child says DEY EE UH NO, he may mean *take me up, no?*, *they eat up, no?*, or final *no* may represent negation as it often does in child syntax, *don't take me up*, or *they are not eating up*.[1] If one attempts the 'equivalence of repetition' and instructs the child to 'Say that

Robert W. Albright (Ph.D., Stanford University, 1953) is Chairman of the Department of Speech and Drama at Arizona State College. Dr. Robert W. and Mrs. Joy Buck Albright (B.A., University of Washington, 1940) have their residence in Tempe, Arizona. This article is adapted from a paper presented at the 1956 convention of the Speech Association of America in Chicago.

[1]In the following discussion, orthographic representation of conventional speech forms will be in italics and those of child forms in small caps; phonetic transcription will be bracketed, and phonemic transcription will be in boldface type.

again,' the response will usually be something such as DEN? (*again?*) or TEY WHA ANEN? (*say what again?*). So one must usually wait until the speech forms recur in the discourse, in order to contrast them in environments that are as identical as possible.

The collecting of data by tape-recording should be done informally in order to elict samples of speech that are as representative as possible of a child's language system. For example, restrictive procedures such as merely naming objects or pictures may elicit utterances which are mere 'formulas' with limited sound environments: DAT'S A (*that's a*), ISA (*it's a*) or DERSA (*there's a*). Objects and pictures are useful in stimulating free, running speech, however, and in providing referents for the speech forms that the child uses. Parents and older siblings are invaluable as 'translators' in the sampling situation and later when the data are being transcribed and analyzed.

After a sample of one to two hours of continuous discourse has been tape-recorded, the recording is transcribed into a 'narrow' phonetic transcription.[2] For greater accuracy, it is probably best to have at least two persons transcribe independently and check with one another at intervals of 25 or 50 utterances, depending upon the complexity of the data. Disagreements may be resolved by rechecking the tape-recording and, if the point at issue seems crucial, by resampling the child-informant. Older siblings and parents who have a good grasp of the child language being studied can be helpful in resolving problems, although there is usually a residue of speech characteristics, sound features, and so forth, which cannot be transcribed with confidence. For example, it is a fairly common mistake in transcribing child English to 'mishear' the glottal stop [ʔ] as one of the conventional stops: [p, t, k]. By having at least two investigators work together, errors of 'mishearing' various features of a child's language are reduced.

The danger of mishearing sounds has caused some linguists to suggest the use of special symbols in order to avoid the associational influences of familiar letters whose sound-values one will tend to 'hear.' However, the gain in rigor in using special, unfamiliar symbols does not make up for the inconvenience of using them and the later confusion they may cause an investigator who tries to recall exactly what sound-values were assigned to them. Besides, phoneticians have long used the same symbols to represent dissimilar sounds in different languages. The sounds of [t] and [d], for example, are made differently in French and English, yet the letters may be used to represent the apical stop consonants in both languages without confusion. The point is to make sure that the symbol represents a sound in a particular language and is not assigned the value that it has in some other language. In the final presentation of child language studies, however, one might very well use small capitals or other special letters to avoid confusing the child's language structure with that of a conventional language.

'Narrow' transcriptions include features of pitch and stress which occur in a language sample. The authors have attempted to establish 'juncture' (*5*), the boundary features of smaller

[2]The manual found most useful by the authors was that of Bloch and Trager (*2*). For a concise summary of field methods in linguistic work, see Swadesh (*7*), and for special symbols and diacritics, see *The Principles of the IPA* (*6*).

units within a stretch of utterance between pauses, but have not been successful. This failure is partly due to the disagreement among linguists as to how such boundaries should be defined and their features described, and partly to the difficulty of establishing regular features of juncture points in samples of child language.

In analyzing a sample of discourse, a stretch of continuous speech between pauses is usually taken as the utterance unit, unless smaller units can be established. Using the method of paired comparison, recurring utterance units are compared with one another; speech sounds and larger forms that seem to contrast in the same environments are isolated. Thus, minimal sound-units and minimal meaning-units (tentative 'phonemes' and 'morphemes') are distinguished. It is possible that the units of child language should be classified in terms that would fit child speech forms more closely. Perhaps they could be set up as 'infantemes,' units of child language having varying forms. Cohen (*4*) has suggested calling each element of communication an 'élémcom,' the unit of gesture a 'gestcom,' of crying and groaning a 'gémcom' and of speech a 'parlcom.' Whatever unit is employed, it is important to avoid the error of classifying the units of a child's language as if they were the same as those of conventional language.

In analyzing units of utterance, environmental features need to be carefully described. The immediate environments of preceding and following sounds in many utterances may be noted to provide information as to the characteristic sequences of sounds. Other features such as the positional occurrence of sounds in utterances, intonational features and reduplication of sounds also need to be described.

In the following discussion conventional symbols of the International Phonetic Alphabet are used. The numbers placed above letters indicate successively higher pitch levels with '1' for the lowest, '3' the highest, and '2' the average pitch level. Four degrees of stress are represented by [ˈ, ˆ, ˋ] for heavy, medium, and light, with very light stress unmarked.

This approach helps to get at the systematic structure underlying the variability of forms in a child's language. The functioning of a child's sound system is usually quite complicated, as the following summary of the alternation of glides in the speech of a 26-months-old child suggests. The sound [w] alternated with [l] initially in utterances, [l] was replaced by zero [ø][5] following [ʊ] and [ɔ], and [l] was replaced by [o] following [ɛ] and [t]. The sound [ɬ] was an alternant of [l] before back vowels and after back consonants. The sound [w] alternated with [r] before [ɪ] and in clusters of a stop plus [r] in segments such as [³twɛ́ɪn¹] corresponding to English *train*. The sound [w] replaced [r] in combinationations of a vowel plus [r] before [aɪ]. In sequences of [aʊ] plus [r] followed by a vowel, [r] was replaced by [ə]. Zero [ø] was an alternant of [r] in clusters of a stop plus [r], and of [r] plus a stop: e.g., [³máɪtøʊfòn¹] corresponding to *microphone*, and [³kɔ́ødəˠ¹] corresponding to *recorder*. Syllabic [ɜˠ] was replaced by [ʊ́] before a stop: e.g.,

[5]The concept of zero [ϕ] is important to child language study, as the regular absence of a sound in its patterning in a sound system may be as much of a signal as the regular presence of a sound or other audible feature.

[²ɑ̀ɪ³bʊ́p¹] corresponding to *I burp*. The sounds [o] and [ʊ] replaced [w] in clusters of a consonant plus [w]. Otherwise [w] had no alternants and thus showed less alternation than [l] and [r]. Also [j] had no alternants. This child's sound-system is more completely discussed elsewhere by Albright and Albright (*1*).

In a study of a one-year-old child with a less developed language system, the phonetic transcription indicated a total of 41 sounds, 27 vowels and 14 consonants. After making a tentative analysis of the phonemes, or 'infantemes,' there was a remainder of seven consonants, five vowels, and two diphthongs, a reduction from 41 to 14 sounds as units in a sound-system. The following chart presents the general structure of the younger child's consonant-system:

	Labial	Apical	Velar	Glottal
Stop	b	d	g	ʔ
Fricative				h
Nasal	m	n		

This sound system may be described as one based mainly upon stop-continuant and oral-nasal contrasts. The voiced-voiceless contrast of conventional English did not seem to be established. There was a good deal of variation of [p] with [b] and of [t] with [d] in the child's speech, however, indicating that the voiced-voiceless contrast was probably developing (*3*).

The vowel-system was reduced to ɪ, ɛ, ɑ, ʊ, plus the central vowel ə, and the two compound vowels ɛɪ and ɑɪ. The general structure of the vowel-system may be charted as follows:

ɪ ʊ
ə
ɛ ɑ

When one considers how resistant most languages are to change, it is difficult to understand how children manage to keep changing their language systems as rapidly as they do. In this struggle between maintaining what Jespersen called a 'little language' for communication with others and, at the same time, developing it through successive stages of increasing complexity, lies what is probably the central problem of speech development in children. Careful descriptive studies should help to clarify this development.

Summary

Techniques of descriptive linguistics may be applied to the analysis of samples of child language if such techniques are adapted to particular problems. In the paired comparison of speech forms that are the 'same' or 'different,' one must usually substitute the 'equivalence of recurrence' for the linguist's 'equivalence of repetition.'

Tape-recordings of several hours of a child's informal discourse probably furnish the most adequate samples for attempting linguistic analysis. In transcribing and developing a 'tentative' or 'suggested' phonemic analysis, it is best to have at least two trained investigators check one another's work.

Since there is considerable variation in the language systems of children, it is important to avoid mistaking variant forms for essential units. In spite of the variation, however, and the fact that the child's language system is developing, it still may be treated at a particular time as a self-contained system—a language in its own right.

References

1. ALBRIGHT, R. W., and ALBRIGHT, J. B., The phonology of a two-year-old child. *Word*, 12, 1956, 382-390.
2. BLOCH, B., and TRAGER, G. L., *Outline of Linguistic Analysis*. Baltimore: Ling. Soc. of Amer., 1942.
3. CHAO, Y. R., The non-uniqueness of phonemic solutions of phonetic systems. In *Readings in Linguistics*, ed. by M. Joos. Wash.: Amer. Counc. of Learned Soc., 1957.
4. COHEN, M., Sur l'étude du language enfantin. *Enfance*, 5, 1952, 181-249.
5. GREENBERG, J. H., The linguistic approach. In *Psycholinguistics: a Survey of Theory and Research Problems*, ed. by C. E. Osgood and T. A. Seboek. Bloomington, Ind.: Ind. Univ. Pubs. in Anthrop. and Ling., 1954.
6. INTERNATIONAL PHONETIC ASSOCIATION. *The Principles of IPA*. London: IPA, 1949.
7. SWADESH, M., A method for phonetic accuracy and speed. *Amer. Anthrop.*, 39, 1937, 728-732.
8. TRAGER, G. L., The field of linguistics. *Studies in Ling.*, 1, 1949, 1-8.
9. WELLS, R., Immediate constituents. In *Readings in Linguistics*, ed. by M. Joos. Wash.: Amer. Counc. of Learned Soc., 1957.

Auditory Skills Of Blinded Individuals Training With Pilot Dogs

John J. O'Neill

Herbert J. Oyer

Donald J. Baker

Various aspects of the hearing ability of the blind have been studied by several investigators. However, these studies have dealt, in the main, with aspects other than the acuity of hearing or mere threshold measurements. In a series of studies at Cornell University (*2*, *6*, *10*), it was found that the perception and avoidance of obstacles by the blind was dependent upon certain auditory dimensions, with pitch discrimination being the most valuable auditory skill involved. Ammons, Worchel and Dallenbach (*1*) studied the effects of hearing losses upon the ability of blindfolded persons to perceive objects in an out-of-doors situation. The results of the study indicated again that auditory cues appeared to be the most reliable, accurate and universal of all of the cues that assisted in the perception of obstacles when visual cues were eliminated. The authors concluded that normal or near-normal hearing is important for the blind in such perception only to the extent that it provides the most reliable perceptual cues.

Only a few studies have been made of the actual hearing acuity of the blind. Seashore (*5*) reported that blind subjects did not differ from sighted subjects in terms of loudness discrimination. Kunz (*4*) employed a watch test in a comparison of the hearing of 38 blind subjects and five sighted subjects. He found that sighted subjects had somewhat better hearing. The first audiometric study of the blind was performed by Hayes (*3*) who found that 19 seeing subjects had on the average more acute hearing than 176 blind subjects.

In a more practical sense a major problem with the blinded is one of developing a means to negotiate in their environment. This usually involves a guide of some sort. The auditory avenue has been used in one such guide device, the flashsound or auditory guidance unit. This sort of device enables the blinded individual to obtain a sound message in the perception of obstacles. Objective as well as subjective evaluations of such aids

John J. O'Neill (Ph.D., Ohio State University, 1951) is Associate Professor of Speech, Ohio State University. Herbert J. Oyer (Ph.D., Ohio State University, 1955) is Assistant Professor of Speech, Ohio State University. Donald J. Baker, (M.A., Ohio State University, 1956) is Research Assistant in the Department of Speech, Ohio State University. This study was supported by funds provided by the Ohio State University Development Fund (Project 5515).

indicate that the loudness, pitch and quality changes that arise from the use of such a device enable the blinded individual to observe obstacles (7, *8*, *9*, *11*). Another guide device is the well-known guide or pilot dog. The purpose of the present study was to determine whether a significant relationship existed between the ability of blinded individuals to use a pilot dog successfully and their hearing acuity, hearing discrimination and sound localization ability. The experimental subjects were drawn from Pilot Dogs, Inc., of Columbus, Ohio, which is one of eight centers in the United States established for the training of pilot dogs for sightless persons. The individuals under study were brought to the center to be trained in the use of a pilot dog.

Procedure

Fifty-three subjects who were sent to Pilot Dogs, Inc., for training in the use of a pilot dog served as experimental subjects before, or just at the beginning, of their training period. Forty of the subjects were males and 13 were females. Each subject was administered a battery of tests which included the following: (1) pure tone threshold tests, (2) sound discrimination tests, and (3) sound localization tests. The various tests were administered during one test session. The test session required approximately two hours. However, test time varied slightly from subject to subject depending upon the individual abilities of the subjects. All tests were administered in a sound-treated testing room. Subjects were observed through a glass viewing port located between the testing room and an adjacent control room.

Tests. The hearing thresholds of the subjects were determined by two test methods. The first method employed a conventional pure tone audiometer and conventional pure tone audiometric test procedures. The second test method utilized the Békésy audiometer and a continuous-frequency testing method. The subjects utilized a hand switch to control the intensity of the stimulus tone. The intensity adjustments were graphically recorded.

The average decibel loss for the frequencies 500 to 2000 cps was computed for each ear. In the instance of the Békésy tracings, threshold points were determined for these same frequencies in the following manner: the mid-point between the lowest and highest threshold tracings within the particular frequency range under test was taken to be the threshold point for that particular frequency.

The B series of the Seashore Tests of Musical Talent was utilized to test six measures which seemed to relate to auditory discrimination. These tests are on phonographic disc and consist of tests of pitch, loudness, rhythm, time, timbre and tonal memory. The subject was seated in the sound-treated testing room eight feet from a Voice of the Theatre loudspeaker. The test materials were presented through a high fidelity recording system at a level of 60 decibels. The various tests utilize a pair comparisons method of testing. The usual test procedure requires that test responses be written. However, the manner of test responses was modified. A signal switch which activated a colored lamp in the control room was used by the subjects. The response involved the indication of a single judgment between two alternatives

and the subject merely signaled his choice. Thus if the first tone was louder, higher in pitch, or different in quality, length or rhythm, the subject would depress the signal switch once. However, if the second tone was to be chosen, the subject depressed the switch twice. The experimenter would then record on the test blank the responses of the subject. Responses to the Test of Tonal Memory were given orally by the subject and written down by the experimenter. The tests were scored in terms of the number of test items correct for each of the subjects. This yielded a set of six scores for each subject.

The test of sound localization included four different types of stimulus materials. These were clicks, spondee words, bursts of white noise and 1000-cycle tone pulses. The noise bursts and tone pulses were one-twentieth of a second in duration. Eight spondee words were employed. They were *hotdog, pancake, therefore, footstool, iceberg, daybreak, jackknife,* and *woodwork.* The recorded click was produced from a recording of the snapping of a plastic clothespin. All test materials were presented at an approximate level of 60 decibels. There was an interval of one-half second between like stimulus materials and a one-second interval between different types of materials.

The localization test was administered to one subject at a time. The subject was seated in the center of a sound-treated room. Eight Utah five-inch loudspeakers (Model SP 5 C) were mounted on the walls of the room four feet above floor level. Three of the speakers were equally spaced to the right of the subject and three to the left. One was placed directly in front of the subject and another directly behind him. All loudspeakers were approximately four feet from the subject's position in the test room. Prior to the actual testing, the subject listened to the stimulus record as it was presented through the loudspeaker. This served to acquaint him with the signals he would be hearing during the test proper.

The points of origin of the sound source were randomized. As the stimulus materials were presented, the subject pointed in the direction from which he thought the stimulus had originated. The tester observed the responses of the subject through the glass window located between the sound chamber and the control room and recorded the responses on an especially prepared answer sheet. The tests were scored in terms of the number of correct responses for each of the types of stimuli presented. Thus for each subject there were four test scores for the test of auditory localization.

Results and Discussion

There were 11 measures for each subject. These were (1) two measures of average decibel loss, one for the left and one for the right ear, (2) two additional measures of average decibel loss, one for each ear obtained from the Békésy test, (3) one pooled measure for the four localization tests and (4) six measures on the six parts of the Seashore Measures of Musical Talent test.

The names of the trainees under study were submitted to the two trainers employed in the Pilot Dog training program for an evaluation of each subject's relative skill in the use of a pilot dog. These ratings were made on a three-point scale. The

TABLE 1. Mean pure tone thresholds for three proficiency categories: above average, average and below average.

Ear	*Above*	*Average*	*Below*
Standard test			
Right Ear	2.45	10.96	21.00
Left Ear	3.09	12.14	17.54
Békésy test			
Right Ear	6.12	14.91	33.30
Left Ear	10.26	18.02	28.67

points on the scale were above average, average and below average.

Hearing Thresholds. In the instance of both the routine pure tone tests and the Békésy test, the average loss for each ear for the frequencies 500, 1000 and 2000 cps served as the basic measures. These measures were grouped to correspond with the ratings above average, average and below average assigned to the subjects by the trainers (See Table 1). The differences among the three proficiency categories with respect to mean decibel loss scores, for left ears and for right ears separately, were evaluated by means of the *t*-test. The mean thresholds, for both the right ear and the left ear comparisons, of the above-average group were significantly better than corresponding thresholds for the average and below-average groups (See Table 2). However, in the instance of both the left and right ear thresholds, there was no significant difference between the average and below-average groups.

Apparently, the above-average group had better hearing (lower thresholds) than the average and below-average groups. Thus, the results of analysis of the pure tone data were in agreement with the results of earlier studies; that is, auditory acuity

e

TABLE 2. Results of the *t*-test for evaluation of differences among proficiency categories: (1) abov average (N=11), (2) average (N=29) and (3) below average (N=13).

Test	*Groups Compared*		
	1 vs. 3 t*	*1 vs. 2* t†	*2 vs. 3* t††
Standard threshold**	2.71	2.04	1.80
Standard threshold***	2.94	2.34	1.20
Békésy threshold **	3.44	1.74	2.92
Békésy threshold***	2.87	1.53	1.89
Localization	2.38	0.67	3.02
Seashore Tests			
Pitch	2.47	.94	1.77
Loudness	1.18	.04	1.13
Rhythm	.98	.15	.92
Time	1.44	.64	1.07
Timbre	1.78	1.92	.10
Tonal Memory	1.63	.30	.99

*$t_{.05}$ (df = 22) = 2.07; $t_{.01}$ (df =22) = 2.82
†$t_{.05}$ (df = 38) = 2.03; $t_{.01}$ (df = 38) = 2.72
††$t_{.05}$ (df = 40) = 2.02; $t_{.01}$ (df = 40) = 3.55
**Average loss for right ear
***Average loss for left ear

TABLE 3. Mean correct scores for B series of Seashore Tests of Musical Talent for three proficiency categories: above average, average and below average.

Test	*Above*	*Average*	*Below*
Pitch	30.45	27.86	22.85
Loudness	35.73	35.57	31.38
Rhythm	20.80	20.50	18.30
Time	32.54	30.68	27.46
Timbre	38.64	31.71	31.31
Tonal Memory	14.78	13.83	10.80

led to better localization performance on the part of blinded subjects.

The threshold measures obtained from the Békésy test (See Table 1) were put to similar statistical test. The results of the analyses (See Table 2) indicate there were no significant differences between the mean thresholds of the above-average and average groups either for the left or right ears. However, the differences between the above-average and below-average groups, for both the left and the right ear, were significant. For the average and below-average groups, the difference between right ear thresholds was significant. The Békésy test results thus agreed with the standard test results except for the fact that the differences between the above-average and average groups on the Békésy test, although in the same direction as for the standard test, were not significant. The inconsistency is probably due, in part, to the screening method utilized in the Békésy test.

Localization. The localization measure was the number of correct identifications of the source of the sound pooled for four types of stimulus materials, with a possible total of 30. The scores were grouped to correspond to the three proficiency categories for purposes of statistical analysis. The mean scores for the above-average group were not significantly different from those of the average group but were significantly higher than those of the below-average group (See Table 2). The localization test thus apparently separated those subjects who did poorly on the training program from those who were most successful in the use of a guide dog.

Seashore Test Results. Only one difference among proficiency groups relative to the various Seashore measures was significant, the difference between the pitch discrimination test results for the below-average and above-average groups (See Tables 2 and 3). One significant difference out of a possible 18 can best be attributed to chance. In all other instances, the differences between groups, in terms of the results of the tests of pitch, loudness, rhythm, time, timbre, and tonal memory, were not significant. The Seashore battery appears not to be a very useful discriminatory tool as it was used in this study.

Summary and Conclusion

A study was made of the hearing acuity and discriminative abilities of 53 blinded subjects who were receiving training with pilot dogs. Tests of hearing included (1) conventional pure tone and Békésy audiometric tests, (2) localization tests and (3) the Seashore Tests of Musical Talent.

The instructors in charge of the training of the subjects in the use of pilot dogs rated each subject in terms of relative skill in the use of pilot dogs. This rating was accomplished in terms of a three-point scale (above-

average, average, and below-average). Comparisons were made between above-average, average, and below-average groups with respect to the various test results.

The results indicated that the subjects who were above average in their use of a pilot dog achieved lower pure tone thresholds (conventional and Békésy audiometric tests) than did the average and below-average groups. The above-average group also exhibited superior performance on the tests of localization. The Seashore Test of Musical Talent (Form B) did not distinguish among the three proficiency categories. The results thus suggest that auditory acuity and also the ability to localize sound are determiners of proficiency in the use of a pilot dog by a blinded person.

Acknowledgments

Professor Thomas B. Anderson, formerly of Ohio State University and now of Miami University, assisted in the early phases of the study. Acknowledgment is also made of the valuable technical assistance provided by William L. Dawson of the Speech and Hearing Clinic staff at Ohio State University.

References

1. Ammons, C. H., Worchel, P., and Dallenbach, K. M., 'Facial vision': the perception of obstacles out of doors by blindfolded and blindfolded-deafened subjects. *Amer. J. Psychol.*, 66, 1953, 519-553.
2. Cotzin, M., and Dallenbach, K. M., 'Facial vision': the role of pitch and loudness in the perception of obstacles by the blind. *Amer. J. Psychol.*, 63, 1950, 485-515.
3. Hayes, S. P., *Contributions to a Psychology of Blindness.* New York: American Foundation For The Blind, 1941.
4. Kunz, M., *The Physiology of the Blind.* Washington: Judd and Detweiler, 1908.
5. Seashore, C. E., and Ling, T. L., The comparative sensitiveness of blind and seeing persons. *Psychol. Rev. Monogr.*, 25, 1918, 148-158.
6. Supa, M., Cotzin, M., and Dallenbach, K. M., 'Facial vision': the perception of obstacles by the blind. *Amer. J. Psychol.*, 57, 1944, 133-183.
7. Twersky, V., An obstacle detecting device for the blind. *Bio. Rev. CCNY*, 9, 1947, 16-21.
8. Twersky, V., Flash sounds and aural constructs for the blind. *Phys. To-day*, 4, No. 3, 1951, 10-16.
9. Twersky, V., On the physical basis of the perception of obstacles by the blind. *Amer. J. Psychol.*, 64, 1951, 409-416.
10. Worchel, P., and Dallenbach, K. M., 'Facial vision': perception of obstacles by the deaf-blind. *Amer. J. Psychol.*, 60, 1947, 502-553.
11. Zahl, P. A., *Blindness: Modern Approaches to the Unseen Environment.* Princeton: Princeton Univ. Press, 1950.

Mental Imagery

Charles F. Diehl

Noel C. England

Many individuals, according to Froeschels (7) and others (*8, 11*), show a decided preference for a particular field of ideation or conception. This preference is defined as the 'imagery type,' a condition or attitude of mind (auditory, visual or kinaesthetic) through which one most easily acquires new experiences and knowledge and most readily recalls memories of facts or words.

Implications of this hypothesis are extensive. Acceptance of a mental imagery variable by psychologists would require its inclusion in all learning theories as a significant factor. Acceptance by educators would imply reevaluation of traditional teaching techniques in terms of individual differences of the students involved.

There seems to be some empirical evidence that an imagery-type hypothesis is tenable. Reading, for example, is a learned process. Yet experienced elementary grade teachers frequently observe that the phonic (auditory) method of teaching reading does not work equally well with all children. Some seem to respond more readily to the whole-word (visual) method. In other instances, according to Fernald (*4*), children classified as 'poor readers' have gained normal reading skill by tracing the outlines of new or difficult words with their fingers as they speak them (kinaesthetic method).

It follows logically then that since speech is a learned process, imagery-type may be a factor in the way speech is learned—or not learned, as the case may be.

The scientific value of any hypothesis is usually directly proportional to the ease with which its tenets can be experimentally evaluated. The imagery-type hypothesis presents many problems in this respect. It is currently impossible to measure or observe with any known instruments the mental processes involved in learning to read, to speak, or to do anything else for that matter. Although Walter (*13*) has reported evidence from his study of brain functioning through electroencephalography which shows clearly three distinct imagery-types, there is probably still much to be done before EEG tracings can be accurately related to specific learning processes. An investigator in this area, lacking appropriate instruments, is, therefore, faced with having to use less objective methods of experimentation—methods which at best yield data collected from per-

Charles F. Diehl (Ph.D., Penn. State, 1949) is Professor of Psychology and Director of the Speech Center, University of Kentucky. Noel C. England (M.A., University of Kentucky, 1953) is a graduate student at Columbia University.

sonal judgments and introspective reports, data of questionable reliability.

Although the concept of imagery-type is old (See Charcot (*3*), Galton (*8*), for example), investigators have apparently been fully aware of the experimental limitations surrounding it. There is, in any event, a dearth of objective experimentation in this area. The little that does exist is largely subjective and controversial. Historically, there have been arguments both for (*1, 3, 8*) and against (*2, 5*) the hypothesis. Contradictory conclusions concerning the effect of training on imagery have also been reported. Galton (*8*) and Gates (*9*) believed that the ability to visualize could be developed with special training while Fernald (*5*) and Foster (*6*) found no evidence to support this contention. Roe (*12*), reporting the results of an intensive study of 64 eminent research specialists in the major sciences, noted differences in the introspective reports of the way these scientists used imagery when thinking. She concluded that it could not be deduced from the data whether possession of appropriate imagery is conducive to choice of vocational field or whether work in a field tends to develop a particular type of imagery.

In spite of the obvious problems related to experimental investigation, it seemed that at least some aspects of basic tenets of the imagery-type hypothesis might be explored with fair objectivity. Keeping in mind the availability of certain types of subjects, the authors formulated the following questions for study: (1) Will a random sample (control group) of college students majoring in areas other than art, music and physical education fall into imagery types suggested by Robbins and Robbins (*11*)? (2) Will three random samples of college students majoring in art, music and physical education show statistically significant differences from a control group with respect to mean imagery dominance scores?

It was necessary initally to establish criteria and define certain terms. The suggestions of Robbins and Robbins (*11*), who report that an individual's imagery type can be determined from his reaction to a word list of concrete objects, were followed. Robbins and Robbins list the following four imagery types into which they believe people can be classified on the basis of their reactions: (1) ear-minded, (2) eye-minded, (3) motor-minded and (4) mixed-minded. 'Mental imagery type,' as used here, is therefore defined as the tendency of some individuals to score significantly higher in one of three specific imagery categories (visual, auditory, motor) or one mixed category in their imagery dominance responses to a selected list of concrete objects.

Procedure

The plan of the experiment involved the presentation of a list of 20 selected words to a random sample (control group) of 74 college students majoring in areas other than art, music and physical education and to three experimental groups of college students (30 art majors, 30 music majors and 30 physical education majors). A 'major' was defined as a college student who was at least a junior and registered with the appropriate department (art, music, physical education) of the University of Kentucky as working toward a degree in that department.

Of the 74 subjects in the control group, 37 were females, 37 males (34

juniors, 34 seniors, six graduate students). The mean age of the control group subjects was 22.1 years. Their academic majors included engineering, law, education, English, sociology, history and social work. Of the 30 subjects in the experimental art group, 15 were females, 15 males (eight juniors, 18 seniors, four graduate students). The mean age of the art group was 21.9 years. Of the 30 subjects in the music group, 15 were females, 15 male (18 juniors, 10 seniors, two graduate students). The mean age of the music group was 21.2 years. Of the 30 subjects in the physical education group, 15 were females, 15 males (18 seniors, 11 juniors, one graduate student). The mean age of the physical education group was 21.1 years. On the basis of sex, mean age (approximately) and college classification (approximately), the three experimental groups were assumed to be matched. Each experimental group was also assumed to be matched proportionately with the control group with respect to sex, age and college classification.

Griffitts (*10*) in an investigation of mental imagery used a battery of seven different tests. He reported that one test which involved ratings on a seven-step scale of the names of 75 concrete objects with regard to clearness of mental imagery correlated highly with the other six tests of imagery in his battery. He concluded that this single test of clearness of mental imagery discriminated among the subjects as well as the entire battery. Since Griffitts' concept of clearness of mental imagery corresponds closely with that employed throughout the present study, a 20-word test was formulated from the words used by Griffitts (*10*) and by Robbins and Robbins (*11*).

The 20 test words were *typewriter, telephone, bicycle, piano, fire, car, ball, electric fan, hammer, tennis, clock, whistle, scissors, music, lawn mower, bell, gun, radio, steam* and *fire engine*. They were selected as meeting the following criteria: (1) they named objects which appeared to be common to the experience of people at any educational level and (2) they named objects which could be expected to have stimulus value in all major imagery categories.

The test words were administered in a pilot study to a group of 20 male college students having a variety of academic interests and ranging in college classification from freshman to senior (five from each). All 20 words were found to elicit imagery ratings in all major categories (visual, auditory, motor). The test, therefore, was considered to have good face validity.

To determine the reliability of the ratings for each imagery classification an odd-even split correlation was computed on the ratings of all words by all subjects for each imagery type. The visual ratings were found to have an odd-even correlation of .88; the auditory ratings, .83; and the motor ratings, .82. Ratings in the 'anything else' category on the dominance scale were too infrequent for statistical treatment.

The test words were presented to 164 subjects. Some subjects completed the test individually, others as part of small groups of three or four. A possible systematic order effect was avoided throughout the experiment by varying both the order and method of word presentation. For half the subjects the first 10 words were presented visually by use of word-cards, manually exposed; the last 10 were presented orally by reading them. For

the other half of the subjects the last 10 words were presented first, visually, then the first 10 orally. No time limit was set between stimulus presentation and response, and the response times were not recorded.

The subjects were asked to report how the objects came into their minds as they thought about them. They were told, for example, that if the word 'pencil sharpener' were given, some might think of it by the 'way one looks' when hanging on the wall where they usually sharpen their pencils. Others might think of it by the 'way one sounds' when being operated. Some might think of themselves as actually turning the handle of a pencil sharpener or being involved in some other 'muscular activity' with the object. Still others might think of the odor of a newly-sharpened pencil.

The subjects were further instructed that the experimenter was interested not only in how the objects came into their minds but also in the dominance of the associations—whether they came singly, for example, 'by the way the object looks,' or more than one way, for example, both 'by the way the object looks' and 'by the way it sounds.'

To report this information the subjects were given a mimeographed, equal-appearing intervals rating scale varying in 'dominance value' from *zero* to *seven* for each category. They were asked to rate the dominance of each word in all four categories: (1) by the way it looks, (2) by the way it sounds, (3) by some muscular activity and (4) by anything else. The *zero* value was to be encircled if nothing came into their minds in a certain category. If the object seemed to have an extremely strong dominance under a certain imagery category, it was to be given a value of *seven*. Since there were 20 stimulus words rated on a *zero* to *seven* scale by each subject for each imagery category, the measure for each category for each subject could vary between *zero* and *140*.

Results

The mean imagery dominance scores in all categories were computed for all groups. For the control group the mean imagery dominance scores were as follows: visual category, 70.14; auditory category, 78.84; motor category, 42.53; and other category, 14.88. The difference between the means of the visual and the auditory categories was 8.70 in favor of the

TABLE 1. Results of *t*-tests for related measures between the means of the imagery categories for the 74 control group subjects.

Comparison Categories	M_1*	M_2	$M_{diff.}$	$SE_{diff.}$	t	df	*Sig.*
Visual vs. Auditory	70.14	78.84	8.70	4.55	1.91	73	.06
Visual vs. Motor	70.14	42.53	27.61	4.43	6.23	73	.001
Visual vs. Other	70.14	14.88	55.26	4.09	13.51	73	.001
Auditory vs. Motor	78.84	42.53	36.31	3.93	9.24	73	.001
Auditory vs. Other	78.84	14.88	63.96	3.55	18.02	73	.001
Motor vs. Other	42.53	14.88	27.65	3.40	8.13	73	.001

*The means of the comparison categories are listed in the same order as the categories are listed.

TABLE 2. Results of *t*-tests for unrelated measures between the means of the control and the means of the art majors for each imagery category.

Imagery Category	*M Control*	*M Art*	$M_{diff.}$	$SE_{diff.}$	*t*	*df*	*Sig.*
Visual	70.14	97.20	27.06	3.80	7.12	102	.001
Auditory	78.84	78.30	.54	5.40	.10	102	—
Motor	42.53	41.77	.76	4.75	.16	102	—
Other	14.88	19.27	4.39	4.26	1.03	102	—

auditory category. The *t*-test result was significant at the .06 level. All other differences between types were significant beyond the .001 level (See Table 1).

For the art group the mean imagery dominance scores were as follows: visual category, 97.20; auditory category, 78.30; motor category, 41.77; and other category, 19.27. The difference between the means of the control group and the art group for the visual category was 27.06 in favor of the art group. This difference is significant beyond the .001 level. Mean differences between the control group and the art majors in the other three categories were not significant (See Table 2).

For the music group the mean imagery dominance scores were as follows: visual category, 60.30; auditory category, 76.87; motor category, 44.70; and other category, 12.40. None of these means was significantly different from the corresponding mean of the control group (See Table 3).

For the physical education group the mean imagery dominance scores were as follows: visual category, 75.10; auditory category, 84.03; motor category, 64.27; and other category, 12.63. The difference between the means of the control group and the physical education group for the motor category was 21.74 in favor of the physical education group. This difference is significant beyond the .001 level. Mean differences between the control group and the physical education majors in the other three categories were not significant (See Table 4).

Discussion and Conclusions

Within the limitations of this study the following conclusions appear warranted: (1) College students majoring in areas other than music, art and

TABLE 3. Results of *t*-tests for unrelated measures between the means of the control group and the means of the music majors for each imagery category.

Imagery Category	*M Control*	*M Music*	$M_{diff.}$	$SE_{diff.}$	*t*	*df*	*Sig.*
Visual	70.14	60.30	9.84	6.69	1.47	102	—
Auditory	78.84	76.87	1.97	4.80	.41	102	—
Motor	42.53	44.70	2.17	5.17	.42	102	—
Other	14.88	12.40	2.48	5.06	.49	102	—

TABLE 4. Results of *t*-tests for unrelated measures between the means of the control group and the means of the physical education majors for each imagery category.

Imagery Category	*M Control*	*M Phys. Ed.*	$M_{diff.}$	$SE_{diff.}$	*t*	*df*	*Sig.*
Visual	70.14	75.10	4.96	3.73	1.33	102	—
Auditory	78.84	84.03	5.19	4.76	1.09	102	—
Motor	42.53	64.27	21.74	4.42	4.92	102	.001
Other	14.88	12.63	2.25	3.52	.64	102	—

physical education fall into several imagery types. The fact that mean scores for auditory-, visual-, motor-, and mixed-mindedness among the subjects differed significantly seems to justify the use of imagery-type categories in describing a part of human behavior. (2) College students majoring in art appear to be more 'visual-minded' than a matched control group composed of students having other academic interests. (3) Physical education majors appear to be more 'motor-minded' than the control group. (4) College music majors do not appear to be significantly different in imagery-type from the control group.

Any conclusions drawn on the basis of data obtained from this investigation must be tentative. Presuming, however, that mental imagery types do exist—and this hypothesis cannot be rejected in view of the evidence reported—there are several direct implications for speech correction. More attention might well be given to the recommendations of Froeschels (7) and Robbins and Robbins (*11*), who suggest that retraining the aphasic patient is greatly facilitated by determining his pre-morbid imagery type and suiting the technique of retraining to it. It would obviously be difficult, if not impossible, to determine experimentally whether the cerebral insult had or had not disturbed the original imagery-type. Those who have worked with aphasic patients, however, generally agree that at least in some cases techniques and materials used in helping the patient recover are determined by the patient—not the therapist. Visual stimuli, for example, will work well for one patient and poorly for another.

Another implication is related to the large category of 'functional articulation cases.' One could postulate for some of them, at least, that (1) the major amount of language stimulation received did not correspond with the dominant imagery-type, (2) a deficiency in one or more major imagery-types exists or (3) the imagery-type is atypical.

Other implications for speech (stuttering, delayed speech, functional voice disorders) and other learned skills (reading, spelling) could be developed from the imagery-type hypothesis. But it is hoped that enough has been said to justify making the following points: (1) The imagery-type hypothesis, if supported by further research, seems to have important theoretical implications for the teaching of speech and other learned skills. (2) The hypothesis cannot, at least on the basis of the evidence reported in this investigation, be rejected. (3) Although investigation of the hypothesis presents experimental problems, some of its tenets can be

experimentally evaluated with fair objectivity. (4) The imagery-type hypothesis is worthy of much more research.

Summary

A control group of 74 college students having academic interests other than art, music, or physical education was compared for imagery-types (visual, auditory, motor, other) with three experimental groups of college students (30 art majors, 30 music majors, 30 physical education majors). All subjects were given ratings of imagery dominance on the basis of their responses to a specially prepared word list representing 20 concrete objects.

The results revealed that for the control group the mean dominance scores differed from one category to another. The mean dominance score for the visual category was larger for art majors than for the control group. The mean dominance score for the motor category was larger for physical education majors than for the control group. No evidence of differences between the music group and the control group was found.

Implications of the data for speech correction were discussed.

References

1. Armstrong Jr., A. C., The imagery of American students. *Psychol. Rev.*, 1, 1894, 496-505.
2. Betts, G. H., *The Distribution and Functions of Mental Imagery*. New York: Teachers College, Columbia Univ. Press, 1909.
3. Charcot, J. M., and Pitres, A., *Clinique des maladies du systeme nerveux*. 2 vols. Paris: F. Alcan, 1893.
4. Fernald, G. M., *Remedial Techniques in Basic School Subjects*. New York: McGraw-Hill, 1943.
5. Fernald, M. R., The diagnosis of mental imagery. *Psychol. Monogr.*, 14, 1912, 1-169.
6. Foster, W. S., The effects of practice upon visualizing and upon the reproduction of visual impressions. *J. educ. Psychol.*, 2, 1911, 11-22.
7. Froeschels, E., and Jellinek, A. *Practice of Voice and Speech Therapy*. Boston: Expression Co., 1941.
8. Galton, F., *Inquiries into Human Faculty and Its Development*. New York: E. P. Dutton, 1908.
9. Gates, A. I., *The Improvement of Reading*. 3rd. ed. New York: McMillan, 1947.
10. Griffitts, C., Individual differences in imagery. *Psychol. Monogr.*, 37, 1927, 1-91.
11. Robbins, S., and Robbins, R., *Speech Sound Discrimination Tests*. Boston: Expression Co., 1948.
12. Roe, A., A study of imagery in research scientists. *J. Personality*, 19, 1951, 459-470.
13. Walter, W. G., *The Living Brain*, New York: W. W. Norton, 1953.

Traditional And Psychogalvanic Skin Response Audiometry

Kenneth W. Burk

Evidence concerning the validity of psychogalvanic skin response (PGSR) audiometry is still considered to be somewhat inadequate by some audiologists. The studies which have been made, however, tend to support PGSR audiometry as a useful clinical method.

Doerfler and McClure (*3*), using a sample of 30 adults with conductive hearing losses at 4000 cps, report a mean difference of 3.50 db between traditional and PGSR thresholds, with the former showing less average hearing loss.

Hardy and Bordley (*4*) report comparisons between traditional and PGSR thresholds for two separate samples of clinical cases in a non-experimental situation. In one sample of 30 ears from adults with known hearing impairment at the octave frequencies 500-4000 cps, the mean deviation was plus or minus 2.46 db. In a second sample of 53 ears of children with known hearing impairment at the same frequencies, a mean deviation of plus or minus 3.57 db was found. The traditional thresholds were consistently better (less hearing loss) than those for PGSR.

This article reports the results of an experiment designed to obtain a further estimate of the validity of PGSR audiometry, using the frequencies 250, 1000 and 4000 cps. The validity criterion was traditional audiometry.

Procedures

All the testing was conducted in a sound-treated laboratory. Data were collected by two experimenters. The PGSR tester had no knowledge of the results obtained by the traditional tester and vice versa. The order of the two techniques was counterbalanced, and all data for each subject were taken consecutively.

Subjects. The 20 subjects used in the experiment were 10 males and 10 females selected randomly from the 363 students enrolled in the introductory course in psychology at the University of Kentucky. Their hearing was expected to fall within the range of normal sensitivity.

The females ranged in age from 17 to 24 years, with a mean age of 18 years, 10 months. The males had an age range of 18 to 27 years, with a

Kenneth W. Burk (M.A., University of Kentucky, 1955) is a speech therapist, Speech Center, Department of Psychology, University of Kentucky. This article is based on a master's thesis completed at the University of Kentucky under the direction of Dr. Frank Kodman, Jr., and Dr. Charles F. Diehl.

mean age of 21 years, 11 months. The mean age for all subjects was 20 years, four months.

Traditional Thresholds. Traditional thresholds were obtained using a recently calibrated Maico, Model MA-1, audiometer. A modification of the Curry and Kurtzrock ear choice technique was employed (*2*) in taking a standard hearing loss audiogram of the octave frequencies 125-8000 cps for each subject. Each threshold was the mean of two ascending and two descending crossings.

PGSR Thresholds. The apparatus for obtaining PGSR thresholds consisted of the Maico, Model MA-1, audiometer, an Esterline-Angus Model A.W. recorder (rate of one-half inch per five seconds), and a locally constructed DC amplifier with a built-in shock component (maximum shock intensity of five milliamperes). Opaque goggles were worn by the subjects to control the influence of extraneous visual stimuli. As a result of the findings of Conklin (*1*), the temperature in the laboratory was maintained at approximately 85° F. by means of an automatic electronic heater.

The common and pickup electrodes were taped to the fingertips of the third and fourth fingers of the left hand. The shock electrode was applied to the neck. Following the reading of standard instructions, the subject was administered a 2.5 milliampere shock with the explanation that his sensitivity to shock was being determined. This tended to sensitize him for the conditioning that followed.

Conditioning was begun with a 250 or 4000 cps tone at a sensation level well above threshold. Following a one-second interval a manually controlled shock of about one-half second duration was administered. The inter-trial time interval was arbitrarily varied between 20 and 40 seconds. The shock level was set at one milliampere and its value altered at the discretion of the tester. The subject used a signal light to indicate when the shock level became too intense.

Threshold explorations were begun when the continuous recorder showed three successive PGSR's of greater than five millimeter amplitude occurring within approximately three seconds after the termination of the test tone. Four thresholds (two frequencies for each ear) were then taken for each subject. The initial frequency was counterbalanced between 250 and 4000 cps. Thresholds were then taken on both ears for a third frequency (1000 cps) for subjects not too fatigued. All the female and half of the male subjects were tested at 1000 cps.

Each threshold was the mean of two ascending and two descending crossings, using five-db steps. The initial sensation level of the tone was varied for the descending series. The order of crossings, as well as which ear was stimulated, was also varied. A response was considered valid if it occurred within approximately three seconds after the termination of the test tone and had a minimum amplitude of two millimeters from the baseline as determined by visual inspection.

Following the findings of Meritser and Doerfler (*5*), 40% reinforcement was employed during threshold determinations. The time relationships were the same as those used during the initial conditioning, and the intensity of the shock was again varied at the discretion of the tester.

TABLE 1. Mean traditional and mean PGSR thresholds at each of three frequencies with result of the *t* test for related measures.

Frequency (cps)	*Test*	*N*	*df*	*Mean (db)*	*Sigma*	*Mean Diff.*	*S. E. Diff.*	*t*
250	PGSR	20	19	1.60	16.28	5.60	4.44	1.26*
	Tra.	20		−4.00	10.91			
1000	PGSR	15	14	5.37	21.67	3.04	7.16	.42*
	Tra.	15		2.33	16.11			
4000	PGSR	20	19	9.23	25.61	1.98	7.37	.27*
	Tra.	20		7.25	19.84			

*Not significant.

Results

The results are based on thresholds taken on only the right ears of the subjects. The data for this ear were selected by chance after a *t* test for related measures revealed that the thresholds for the left and right ears did not differ significantly. Thresholds were obtained for all 20 subjects at 250 and 4000 cps and for 15 subjects, 10 female and five male, at 1000 cps.

The mean difference between traditional and PGSR thresholds at each frequency tested (250, 1000 and 4000 cps) was evaluated by the *t* test for related measures. None of these differences was statistically significant (See Table 1). Inter-frequency comparisons were also made using the *t* test for related measures. None of these differences was significant at the five per cent level.

Within the limitations of the sample and by comparisons to thresholds obtained by traditional methods, the data appear to support the validity of PGSR audiometry for thresholds taken at 250, 1000 and 4000 cps. The mean test difference at each frequency is comparable to differences reported by other investigators. The results also indicate that this clinical method is consistent in its measurement over all three frequencies.

Summary

Twenty subjects (10 male, 10 female) were selected randomly from 363 introductory psychology students. Thresholds taken on the right ears at 250, 1000 and 4000 cps by traditional and PGSR audiometry were employed to evaluate the validity of the latter technique. Only 15 subjects were tested at 1000 cps.

The PGSR audiometric method was found to give valid measures of auditory thresholds at all three frequencies

with results of the traditional test as the criterion. PGSR audiometry also showed consistent measurements over all three frequencies.

References

1. Conklin, J. E., Three factors affecting the general level of electrical skin-resistance. *Amer. J. Psychol.*, 64, 1951, 78-86.
2. Curry, E. T., and Kurtzrock, G. H., A preliminary investigation of the early-choice technique in threshold audiometry. *JSHD*, 16, 1951, 340-345.
3. Doerfler, L. G., and McClure, C. T., The measurement of hearing loss in adults by galvanic skin response. *JSHD*, 19, 1954, 184-189.
4. Hardy, W. G., and Bordley, J. E., Evaluation of hearing in young children. *Acta Otolaryng., Stockh.*, 40, 1951-2, 346-360.
5. Meritser, C. L., and Doerfler, L. G., The conditioned galvanic skin response under two modes of reinforcement. *JSHD*, 19, 1954, 350-359.

Language Skills In Children With Cleft Palates

D. C. Spriestersbach

Frederic L. Darley

Hughlett L. Morris

The question of language retardation in children with cleft lips and palates has been considered by several writers. Brown (*4*) lists poor physical development resulting from feeding problems, repeated illnesses, difficulty with hearing, isolation during illnesses or hospital confinements, and over-protection by parents as factors that might possibly result in delayed language development in such children. McWilliams (*6*) discusses the importance of intelligence, hearing, physical development and emotional growth as factors influencing the development of language in children with cleft palates.

Spriestersbach, Darley and Rouse (7) have reported that cleft palate children are generally retarded in articulation skills. No published reports are to be found of research designed expressly for the purpose of investigating the *language status* of cleft palate children. The reliability and validity of certain measures of language development, however, have been established by McCarthy (*5*), Davis (*3*), Williams *et al.* (*10*) and Templin (*8*).

In the present study the hypothesis tested was that children with cleft palates are retarded in their language development. The procedure consisted of (1) securing measurements of the performance of a group of children with cleft palates using three indices of language development and (2) comparing the performance of the children according to these measures of language development with norms established by Templin (*8*) and Ammons and his associates (*1*, *2*).

Procedure

Subjects. Forty Caucasian singletons (children of single births with or without siblings), 19 girls and 21 boys, born with clefts of the palate were selected for this study. (Twins were excluded so that the language performance of the group could be compared with Templin's norms, which were based upon a study of

D. C. Spriestersbach (Ph.D., University of Iowa, 1948) is Professor of Speech Pathology and Audiology, University of Iowa. Frederic L. Darley (Ph.D., University of Iowa, 1950) is Associate Professor of Speech Pathology and Audiology, University of Iowa. Hughlett L. Morris (M.A., University of Iowa, 1957) is a half-time Research Fellow, University of Iowa. This investigation was supported in part by research grant M-1158 from the National Institute of Mental Health, Public Health Service.

480 singletons.) Twenty-eight of the children had clefts of both the lip and palate while 12 had clefts of the palate only. The subjects ranged in age from from three years and six months to eight years and five months. They ranged in I.Q. or I.Q.-equivalent scores[1] from 80 to 138 with a mean of 101.4. None of the children had an educationally handicapping hearing loss.[2] Their speech was judged sufficiently intelligible to allow for accurate scoring of 50 responses. The distribution of the subjects according to the classification of their fathers' occupations is not identical to that of the population sampled by Templin (*8*), the major difference being that the fathers of 37.5% of this group were engaged in farming whereas no fathers of children studied by Templin were farmers. Comparison of parental occupations of this group with those of the groups studied by Ammons (*1, 2*) is not possible since different classification systems were used. The sample used in this study, however, seemed to resemble, in general, the populations sampled by Templin and Ammons closely enough to permit comparison of their performances on certain language skills.

Measures of Language Development. The following measures were employed: (1) mean length of response, (2) structural complexity score (both of these as determined by analysis of 50 responses in connected speech), and (3) vocabulary size, as measured by the Ammons Full-Range Picture Vocabulary Test. Two factors were considered in selecting these measures. Other investigators have used all three measures in previous studies and have reported coefficients of reliability and validity of sufficient magnitude to warrant their use as research instruments. In addition, norms were available for the three measures with which the performance of this group of children could be compared.

Fifty responses were elicited and recorded and the mean length of the 50 responses was computed in a manner described by McCarthy (*5*). Instructions presented by McCarthy and Davis (*3*) were used in counting the words.

The analysis of language structure was performed by first using the classification outline devised by McCarthy (*5*) and modified by Davis (*3*) and placing each response in one of five categories set up according to completeness and complexity. The responses were then quantified with regard to their structural complexity by means of the Templin (*8*) modification of Williams' (*10*) method. A value from *zero* to *four* was assigned each response, according to its completeness and complexity as determined by the McCarthy-Davis outline. The obtained measures for the entire 50 responses were then added to determine the child's structural complexity score.

Form A of the Ammons Full-Range

[1]Twenty-six children were administered the Wechsler Intelligence Scale for Children. Fourteen children were given the Ammons Full-Range Picture Vocabulary Test and I. Q.-equivalent scores were computed.

[2]An educationally handicapping hearing loss was defined as a loss in the better ear of 30 db or more in one of the frequencies tested (250, 500, 1000, 2000, and 4000 cps) or of 20 db in three of the five indicated frequencies. If a child could not be tested by pure tone audiometry, speech audiometry was used. In such instances a hearing loss was defined as a loss of 20 db above the speech reception threshold for normal *young adults*.

Picture Vocabulary Test was administered in the manner prescribed by the authors of the test (*1*).

Reliability of Recording and Scoring. To determine the reliability of the method of recording the responses, the experimenter and one observer independently recorded the same 50 responses for three subjects chosen at random. All records were then scored by the experimenter and the scores for mean length and structural complexity were compared. Differences between these scores were small and were judged to indicate that the method of recording was sufficiently reliable.

To determine the reliability of the method of scoring the responses, the experimenter and one observer independently scored three of the connected speech samples recorded by the experimenter. Again very small differences indicated that the method of scoring the responses was sufficiently reliable.

Results

Stability of Language Measures. Odd-even reliability coefficients for mean length of response and structural complexity were obtained in order to determine how consistently the children used responses in the various categories of the structural classification system. The mean lengths for even-numbered responses were correlated with the mean lengths for the odd-numbered responses for 32 children. A correlation of .95 was obtained. This finding is higher than the reliability coefficient of .91 reported for mean length by McCarthy (*5*) or of .86 reported for mean length by Williams (*10*). It appears, then, that this group of children were highly consistent in the length of their responses during the elicitation of the speech sample.

The means of the structural complexity scores for even-numbered responses were correlated with the means of the structural complexity scores for odd-numbered responses for 32 children. A correlation of .80 was obtained. This finding is not directly comparable to findings reported by other experimenters. Templin (*8*) is the only experimenter who has used the particular technique employed in this study and she does not report reliability information. McCarthy (*5*) reports reliability coefficients of from .59 to .92 for the various categories in her classification outline. Williams (*10*), using a weighting system similar to that employed in this study, reports reliability coefficients of .89 for sentence completeness and .87 for sentence complexity. This group of children thus appear to be relatively consistent in the structural complexity of their responses.

Description of Language Development. The mean lengths of response used by the subjects were compared with the mean lengths obtained from children of comparable ages studied by Templin (*8*). The differences between the mean lengths of response of the subjects and the corresponding norms ranged from 3.8 words longer than the norm for the subject with the highest score to 5.2 words shorter than the norm for the subject with the lowest score. A mean difference of 1.01 words and a median difference of 1.2 words shorter than the responses of the normative groups were obtained. The standard deviation of the differences was 1.74 words. The

difference between the two groups, was found to be significant beyond the five per cent level ($t = 3.62$; for 39 df, $t_{.05} = 2.02$).

In order to determine the status of individual subjects in terms of number of years of advancement or retardation in mean length of response, the children's ages were compared with the normative age levels corresponding to the scores for mean length that their scores most closely resembled. The difference between the two age levels was determined. Two problems were encountered in making this comparison. The first was the handling of reversals in the norms, that is, a lower normative score for a given age level than for the level immediately preceding it. An example of such a reversal is the norm of 5.4 words for mean length of response at four and one-half years followed by a norm of 5.3 words listed for age five. In such instances the higher of the two age levels was used. In the case cited, a score of 5.4 would be selected as corresponding to an age level of five. Another problem was the handling of norm limits. Since Templin (*8*) did not supply norms for age levels above eight years or below three years, it was impossible to evaluate performances beyond these limits. In instances such as these, the age level corresponding to the norm that best matched the score was used. On the basis of this comparison, five subjects were found to be advanced in mean length of response, six were at appropriate age levels, and 29 were retarded. The majority of the experimental group, then, used responses that were shorter in length than the responses used by Templin's normative group.

Structural complexity scores were compared with the norms reported by Templin (*8*). Differences in complexity scores ranged from +73.1 for the subject with the largest complexity score to −58.2 for the subject with the smallest complexity score. The mean difference between the obtained scores and corresponding norms was −4.77 with a standard deviation of the differences of 26.9. The median difference was −3.6. The subjects were thus on the average retarded in language complexity. However, the mean difference between the scores of the subjects and the norms was not significant at the five per cent level ($t = 1.11$; for 39 df, $t_{.05} = 2.02$).

The subjects' ages were compared with age levels corresponding to the norms that their structural complexity scores most closely resembled. Problems in dealing with reversals in the normative scores and norm limits were handled in the same way as in the case of mean length of response. Structural complexity scores superior to the norms were obtained for 13 subjects, at the norms for three subjects, and below the norms for 24 subjects. Thus more children were retarded in structural complexity scores than were not, but, considered as a group, these children used sentences with a structural complexity that closely resembled that of sentences used by the normative group.

The children's performances on the Ammons Full-Range Picture Vocabulary Test were compared with norms reported by Ammons and his associates (*1*, *2*). The range was from 16.1 more words to 10.3 fewer words for the experimental group than for the normative group. A mean difference of 2.37 words in favor of the experimental group was obtained with a standard deviation of the differences of 5.03. The median difference was

2.4 words. The mean difference between the scores of the two groups was significant beyond the five per cent level (t=2.94; for 39 df, $t_{.05}$= 2.02).

The children's ages were compared with age levels corresponding to the norms that their vocabulary scores most closely resembled and the amount of advancement or retardation was computed. The problem of norm limits was not present since norms have been reported by Ammons *et al.* (*1*, *2*) for levels from two through 16 years of age. Thirteen subjects showed superior performance on the vocabulary test, 10 subjects performed approximately at the appropriate age levels and 17 subjects were retarded.

Discussion

That this group of children with cleft palates was somewhat retarded in mean length of response is not surprising, since it may be that they do not find speech as pleasant, as satisfying, or as rewarding as do children without clefts. Many of them may even have been penalized for their speech in given situations, and, as a probable consequence, talk less than other children of comparable age.

The subjects were not significantly different from the norms in the structural complexity of their remarks. This finding may indicate that any unpleasantness they might have experienced in talking has not affected their syntactical proficiency.

This group of subjects showed superior performances on the Ammons Full-Range Picture Vocabulary Test. In light of a relatively high correlation between vocabulary size and intelligence, this might have been an expected finding had the group been of higher than average intelligence. However, since the mean of the I. Q.'s and I. Q.-equivalents for the group was found to be 101.4, this explanation is not tenable. The probability of errors in the norms seems rather unlikely since Templin (*8*) reports findings comparable to those reported by Ammons (*2*) for the three- to five-year-old group when she grouped her subjects at whole-year levels. (However, when she grouped her subjects at half-year levels, she found higher means than those reported by Ammons for the three- and four-year-old levels, and lower means at the five-year-old level.)

One source of error in this study may have been in the administration of the Ammons test. The test instructions for managing the problem of guessing are not clear. It is possible that the experimenter was too lax in rechecking possible guesses, thus allowing the subjects to achieve spuriously high scores. Since the Ammons Full-Range Picture Vocabulary Test is a popular clinical tool, more extensive investigation of the test with cleft palate subjects appears to be indicated.

It is to be noted that the Ammons test is an index of the size of recognition vocabulary, since the subject is required simply to point to the picture that best matches the stimulus word. Other vocabulary tests, such as the vocabulary subtest of the Wechsler Intelligence Scale for Children (WISC), measure the subject's ability to define the stimulus word orally. Since scores on the WISC were available for those subjects in the experimental group aged five years and older, it was felt that examination of these vocabulary subtest scores would be of interest here.

The subtest scores earned by 22 subjects in the experimental group were compared with the test age-equivalent scores for the subtest reported for the standardization group by Wechsler (*9*). The method of scoring the subtest requires that each word-item be given a value of two, one, or zero points. The obtained scores were compared with the test age-equivalent scores reported by Wechsler. The range of differences between obtained subtest scores and the test age-equivalent scores was from −15.5 points, the difference for the most retarded subject of the experimental group, to +5.5, the difference for the most advanced subject of the experimental group. Nineteen subjects were retarded in vocabulary size and three were advanced. A mean difference of 4.16 points and a median difference of 6.1 points were both in favor of the normative group. The standard deviation of the differences was 4.36. The mean difference was significant beyond the five per cent level ($t = 4.34$; for 21 df, $t_{.05} = 2.07$). Thus it is seen that while the children of the experimental group performed very well on a recognition vocabulary test, those who were given a vocabulary definition test performed less well on this test.

Cleft palate children of normal intelligence apparently may, in general, be expected to perform less well than their peers in amount of verbal output (measured in this study by mean length of response). Even though a child may be as intelligent, have as large a recognition vocabulary and have as adequate a mastery of the subtleties of grammatical and syntactical expression as his peers, he may still be somewhat inhibited in the extent of his verbal expression because of his self-evaluations, his evaluations of his speech, and his past experiences in unrewarding speaking situations. Such a pattern of performance corresponds roughly to the results reported on the three measures employed in this study.

Summary

Three measures of language development were obtained for 40 children with cleft lips and cleft palates or with cleft palates only. The sample was restricted to children between the ages of three and eight who were singletons of the Caucasian race with normal mental ability and without a handicapping hearing loss.

Fifty responses of connected speech for each child were analyzed for mean length of response and structural complexity and resultant scores were compared with norms for these measures. A recognition vocabulary test was administered to all subjects. Comparison of the resulting scores with established norms indicated that cleft palate children have, on the average, superior recognition vocabularies. Since the results of this comparison might have been unduly influenced by some unknown factor, 22 of the subjects were further tested on vocabulary usage.

The subjects demonstrated no general language retardation but were retarded on measures of amount of verbal output and vocabulary usage. The following conclusions appear warranted: (1) Children with cleft palates are in general retarded in mean length of response. (2) Their language development as measured by structural complexity scores is not different on the average from that of the normative group. (3) They are in general retarded in vocabulary usage.

References

1. AMMONS, R. B., ARNOLD, P. R., and HERRMANN, R. S., The full-range picture vocabulary test; IV. results for a white school population, *J. clin. Psychol.*, 6, 1950, 164-169.
2. AMMONS, R. B., and HOLMES, J. C., The full-range picture vocabulary tests; III. results for a preschool-age population, *Child Developm.*, 20, 1949, 5-14.
3. DAVIS, E. A., *The Development of Linguistic Skill in Twins, Singletons with Siblings, and Only Children from Age Five to Ten Years. (Inst. of Child Welfare Monogr. Ser.*, No. 14), Minneapolis: Univ. of Minn. Press, 1937.
4. JOHNSON, W., BROWN, S. F., CURTIS, J. F., EDNEY, C. W., and KEASTER, J., *Speech Handicapped School Children* (Rev. ed.). New York: Harper, 1956.
5. MCCARTHY, D. A., *The Language Development of the Preschool Child. (Inst. of Child Welfare Monogr. Ser.*, No. 4), Minneapolis: Univ. of Minn. Press, 1930.
6. MCWILLIAMS, B. J., Some observations of environmental factors in speech development of children with cleft palate. *Cleft Palate Bull.*, 6, No. 1, 1956, 4-5.
7. SPRIESTERSBACH, D. C., DARLEY, F. L., and ROUSE, V. V., Articulation of a group of children with cleft lips and palates. *JSHD*, 21, 1956, 436-445.
8. TEMPLIN, M. C., *Certain Language Skills in Children, Their Development and Interrelationships. (Inst. of Child Welfare Monogr. Ser.*, No. 26), Minneapolis: Univ. of Minn. Press, 1957.
9. WECHSLER, D., *Manual for Wechsler Intelligence Scale for Children.* New York: Psycholog. Corp., 1949.
10. WILLIAMS, H. M., MCFARLAND, M. L., and LITTLE, M. F., Development of language and vocabulary in young children. *Univ. Ia. Stud. in Child Welf.*, 13, No. 2, 1937.

Skin Resistance Changes Accompanying The Sidetone Test For Auditory Malingering

Clair N. Hanley

William R. Tiffany

Jacqueline M. Brungard

Since the advent of World War II a great deal of attention has been directed towards investigating the reality of hearing deficits in people suspected of malingering and psychogenic deafness. A search of the literature on auditory testing reveals an abundance of so-called malingering tests, including the well-known Doerfler-Stewart (2), Lombard (7), and Stenger (7) tests, as well as numerous references to many of the clinical signs indicative of feigned deafness, such as unreliability of repeated audiograms, and discrepancies between pure tone and speech hearing tests.

The basic limitations of many of these tests appear to be lack of reliability, dependence upon the patient's naïveté, necessity of expert interpretation, and inability to approximate the 'real' auditory threshold (lack of validity). In view of these limitations, it would appear that some finer test for the determination of 'simulation deafness' is needed—one that is simple to administer, demonstrable to a jury, and capable of reliable results, as well as one that will yield a close approximation of the exact auditory threshold when malingering or, possibly, psychogenic deafness has been determined.

One of the most recent tests which has been used to overcome some of the difficulties encountered in malingering testing is the delayed sidetone test. This test has the advantage of not appearing to be an auditory test at all, inasmuch as the subject is simply asked to read a passage into a microphone while wearing a pair of earphones. As he reads, his voice is recorded and fed back to his earphones at a controlled and known intensity level and with a time delay of from about .1 to .2 seconds. The resultant effect is a change in the subject's speech which he seems largely unable to control.

The following changes have been observed in the behavior of persons

Clair N. Hanley (Ph.D., University of Iowa, 1952) is Associate Professor of Speech, Speech and Hearing Clinic, University of Washington. William R. Tiffany (Ph.D., University of Iowa, 1951) is Associate Professor of Speech, Director of Voice Science Laboratory, University of Washington. Jacqueline M. Brungard (M.A., University of Washington, 1956) is Chief Audiologist, Good Hope Memorial Clinic, Los Angeles, California.

reading under delayed sidetone: (1) pitch control is disturbed; (2) speech becomes slurred; (3) rate is altered from the norm; (4) vocal intensity varies markedly from the usual level, particularly if the feedback intensity level is high and an effect is generated like that observed in the Lombard test; (5) stuttering-like speech occurs, with frequent repetitions, distortions, and omissions of speech sounds and syllables; and (6) the subject fidgets, flushes, squirms and generally gives evidence that he is struggling with a difficult task. Not all of these symptoms are equally observable or equally often observed, of course, but some symptoms are nearly always noticed.

Measurements of the disturbances produced by the delayed sidetone have been, to this time, largely dependent upon subjective judgements of vocal disturbance (4), reduction in rate of reading (3), variations in vocal intensity (5), and the tabulation of reading errors (1). Hanley and Tiffany (3), who found that reading rate reduction is a positive function of intensity of sidetone signal, have noted that detection procedures should employ measures of rate change as well as judgements by expert listeners. Further, they call attention to the wide individual differences in response to the sidetone signals. They observed speech breakdown in some subjects at sensation levels as low as 10 to 20 db, whereas some subjects appeared relatively undisturbed at high sensation levels.

Inasmuch as marked individual differences in resistance to speech breakdown and reduction in reading rate do occur, even at high intensities, the need for further refinement of the delayed sidetone testing technique is obvious. There is a strong need for a more sensitive measure of the effect, one that would not only indicate the reality of the hearing loss, but one that would also afford a more definitive measure of the patient's potential auditory threshold. The present study grew out of the attempt to define such a measure.

The current technique employed by the authors for the evaluation of the delayed feedback effect calls for the subject, wearing headphones, to read a selected passage into a microphone of a tape recorder having separate record and playback circuits. After the calculation of the normal, no-feedback reading rate, intensity of delayed sidetone is raised during successive readings of a test passage, usually in 10 db steps, to a point where speech disturbance is obvious, even to a wholly naïve observer. Statistical comparisons between the normal rates of reading and the rates at successively more intense feedback levels provide the basis for a judgement of the 'true' threshold, supported by the obvious speech disturbance at the higher levels.

During the search for refinement of the above technique the authors have experimented in two directions: toward the construction of speech and non-speech oral tasks which would be more sensitive to the delayed feedback disturbance, and toward the employment of a more sensitive measure of the speech disturbance itself. The most promising clues were furnished by comments of subjects in previous experimentation who had described the sidetone task as 'hard work,' 'startling,' 'unpleasant,' 'nerve-racking' and 'annoying.' These remarks led the authors to assume that the reading task under delayed sidetone produced a good deal of stress for the subject which could, it was believed, be measured objectively.

In connection with the above com-

ments concerning the type of response to delayed sidetone, Woodworth (*8*) has noted that, '. . . in momentary thrills of fear, surprise, embarrassment or expectancy, the sweat glands are stimulated to a slight degree by their nerves, and the result is a momentary change in the electrical condition of the skin, which can be registered by a galvanometer and is called the *psychogalvanic reflex*.' Woodworth (*9*) further notes that, in an experiment of Abel, observers' reports indicated 'that PGR is an index not so much of feeling or emotion as of an attitude of the organism directed toward *overcoming difficulty*.'

In a pilot study involving 30 members of the University of Washington faculty and students, it was noted that there tended to be significant skin resistance (PGR or GSR) changes when sidetone levels were systematically varied. Subjects who experienced delayed sidetone at sensation levels of 80 db showed significant and marked reductions in skin resistance, whereas at lower levels these same alterations were not so consistently noted, nor so dramatic in degree. Thus the hypothesis was advanced that the presentation of delayed sidetone to a hearing subject would cause significant changes in his skin resistance and, further, that the consistency of these changes would be a function of the sensation level of the delayed feedback. Finally, it was hypothesized that such changes would make a simple and workable measure for the detection of feedback response in relatively sophisticated subjects and would constitute a good test of malingering.

Method

The subjects from whom GSR measures were to be made while they read under delayed feedback conditions were 50 college students, 17 females and 33 males. They were randomly assigned to five groups of 10 students each, to receive feedback at the levels 80, 60, 40, 20, and 0 db above each student's threshold for spondee words.

Subjects were seated in a sound-isolated room in which a recording microphone, earphones, GSR pickup electrodes, reading stand, signal light and monitor-speaker were located. In the adjoining control room the observers operated the feedback delay mechanism, GSR and recording equipment. The delayed feedback equipment consisted of a tape recorder capable of producing a .14 second delay and suitable amplifying and calibrated attenuation equipment. The GSR measuring equipment consisted of a *Maico Psychometer*, DC amplifier, and graphic recorder. A special switching arrangement was provided which automatically fed the sidetone to the subject and at the same time marked the margin of the recording chart paper so that the skin resistance fluctuations could be correlated with the presence of the delayed sidetone.

The procedure of the experiment was, briefly, as follows: The relatively naïve subject was placed in the sound-proof room after familiarizing himself with a list of spondee words. The GSR electrodes were fastened to the palmar surface of his non-dominant hand, the earphones were placed on his head, and his speech reception threshold was established. The following instructions were then read:

> I want to find out how people talk when they hear their own voices over earphones, while they are reading. Therefore, I would like you to read this passage into the microphone. During your reading, you should try not to let what you hear in the earphones bother you.

That is, try not to laugh or comment, but attempt to read the passage as normally as possible. The metal plate that I strapped to your hand will be used to measure how relaxed you are during the test. Since this plate is connected to a very sensitive recording device, I want you to hold your hand like this (fingers over the edge of arm desk), and keep it as still as possible throughout the test. Now start reading aloud.

The subject was then given a simple prose passage to read which was long enough to last for the duration of the experiment. For the first minute of his reading (approximately) psychogalvanometer sensitivity adjustments were made. At the end of that minute, when it was assumed that he had 'settled down' to his reading task, the random recorder needle fluctuations were, on the average, about one chart division in extent. Following these adjustments 10 eight-second bursts of delayed sidetone were administered at one of the five experimental levels (randomly assigned).

In order to prevent bias in the judgement of the records, to provide for keying of the sidetone during an interval of relatively constant GSR state, and to provide for a random order of stimulus presentation, two experimenters were used. One experimenter presented the sidetone signals at from 15 to 60 second intervals, in a random manner, dependent upon the quiescence of the GSR response. The second experimenter independently determined the level of feedback and blanked out half of the 20 presentations in an order randomly determined and unknown to the first. In this way it was felt there would be a minimum of preconceived bias both on the part of the experimenter and of the judges who would later evaluate the graphic records of skin resistance change.

Analysis of Graphic Records

The graphic records of skin resistance changes obtained from the 50 subjects in this study were analyzed in two ways. The first type of analysis was largely objective in nature. During the testing it had appeared that the most reliable cue for the determination of the presence or absence of the sidetone (particularly in the records from low-intensity-level presentations, where GSR amplitudes were small) was that of consistency in the GSR latency. If the GSR recorder needle swings, following the 10 sidetone presentations, showed highly consistent latencies, then hearing could be inferred. If latencies were inconsistent, hearing was presumed to be absent or the sidetone intensity was found to be at a lower sensation level. High levels of feedback intensity seemed to be associated, of course, with consistent latencies and high graph amplitudes.

For all sidetone presentations (10 per subject) a measure was made from the graphic records of the response latencies following the onset of the delayed sidetone presentations. Arbitrarily, all graphic peaks of greater magnitude than two chart divisions[1] (about 5 mm of arc in the recorder needle swing) were recorded from one second following the actuation of the delayed sidetone and for a duration of 7.5 seconds. For most of these intervals of measurement only one peak of skin resistance change was noticed and never more than two. Of the total

[1]The chart division is, of course, an arbitrary but nonetheless practical and handy unit for clinical reference. In this study no attempt was made to calibrate in terms of absolute measures of skin resistance change. The graphic recorder used was an *Esterline Angus*, model AWDC Milliammeter 4782.

of 500 intervals of time, 81 were found to have two such peaks. One of these was chosen by the first experimenter on a basis of consistency with other measured response latencies and relative response amplitudes.

The 10 measures for each subject which were obtained in this way were taken as the latency measures of the GSR responses to the sidetone. The mean of the 10 measures for each subject was termed the *mean latency*. This mean latency was established for each individual and means of these mean latencies were calculated for each feedback sensation level.

Two measures of variation were obtained for the latency measures. The first was simply a measure of variation among subjects. This was the variance of the mean latency measures about the mean of the group of 10 subjects. The second and most important of the measures was the measure termed the *mean latency variance*. This was a measure of variation within subjects and was based upon the variation of the 10 latencies for any one subject about the mean for the same subject. It was this measure which was hypothesized to be the best indicator of consistency of latency and, thus, of the presence of delayed sidetone response.

The second type of analysis was a partly subjective one where two 'experts' independently examined the unlabeled and unidentified graphic records of skin resistance changes for the 50 subjects and made judgements as to whether the subjects had received a 'disturbing' auditory stimulus and whether it was at 80, 60, 40, 20, or 0 db sensation level. These judges (Hanley and Tiffany) had no idea to whom the graphs belonged. They did know from the margin keying exactly at which point the sidetone presentations began and ended. They did have, additionally, the latency measures described above. Their judgements, however, were based not only upon latencies but also upon magnitudes of response and any other aspects of the graphic pattern they wished to utilize.

TABLE 1. Mean response latencies and variances for 10 subjects expressed in seconds.

Sensation Level (db)	*Mean Latency*	*Among-subjects Variance*	*Average Within-subjects Variance*
80	3.37	.52	.48
60	3.25	.54	.64
40	3.54	.90	1.45
20	3.30	1.52	2.65
0	3.58	3.25	5.20

Results and Discussion

Table 1 presents the data for the latency means and variances, expressed in seconds, for the five sensation levels employed in the experiment. An examination of this table will at once reveal two obvious facts concerning the data: (1) mean response latencies remain about the same regardless of the level of sidetone employed; and (2) latency variances are inversely related to sensation level.

Particularly strong relationships are shown in Table 1 between sensation level and within-subject, or mean latency variance. The range of variance here is from less than one second at 80 db to more than five seconds at threshold, with a marked rise in variance between 20 db and threshold. Subjects also varied among themselves with respect to response latency and this variation was associated with sensation level in the same way. The table shows a steady increase of

TABLE 2. Number of records judged by 'experts' as resulting from several feedback levels and sensation levels (N = 20).

Sensation Level (db)	*Feedback Levels (db)*				
	80	*60*	*40*	*20*	*0*
80	11	8	1	—	—
60	4	8	6	2	—
40	—	6	7	3	4
20	—	1	1	2	16
0	—	—	—	—	20

variance as sensation level drops, with a marked increase from 20 db to threshold.

Table 2 presents the results of the judgements made by the panel of two 'experts' who, with only the latency measures and the graphs themselves, made judgements as to whether or not feedback had been heard by the subject and at what sensation level. This table shows the number of judgements made which fell into each of the sensation level groups. For example, of the 20 judgements made of the graphs of people receiving 80 db of delayed feedback, 11 were judged to have received 80 db, eight were judged to have received 60 db, and one was judged to have received 40 db of feedback. None was judged to have received feedback at levels lower than 40 db.

Of the 100 judgements made by the two observers only 20 were in error to the extent that listeners who actually 'heard' feedback were judged not to have been influenced by it to any extent. Of these, 16 received feedback at only 20 db and four at 40 db above spondee threshold. More important is the fact that of the 100 judgements made by the two observers only eight varied by more than 20 db from the actual sensation level administered. Clinically, this finding tends to indicate the possibility that in over 90% of the cases a patient's threshold may be determined within 20 db by the method employed here.

The above findings would not mean very much if it could be shown that judgements of feedback effect were so liberal as to produce the error of judging as disturbed those who had not heard the feedback. However, it is to be noted that in *not one* case was a subject 'falsely accused.' All 10 of the graphs for subjects receiving feedback at 0 db were judged correctly. This is particularly important for the evaluation of this test as a detector of malingering, and is better than the results obtained by Tiffany and Hanley (*6*) with rate measures alone. Thus, viewed solely in its application to the task of detecting the auditory malingerer, the GSR measure appears to be a valuable supplement to the rate measure and perhaps to be superior to it in some respects.

In addition to demonstrating the reliability and usefulness of a GSR measure of the response to delayed sidetone, the present study has raised some interesting questions concerning the use of the delayed feedback as a sufficient unconditioned stimulus in hearing testing for other purposes. The data show responses of excellent consistency with no previous conditioning and under the simplest of clinical conditions. The reliability of this method of GSR testing should be studied further with a view toward developing 'disturbances' of sounds or of speech-hearing processes as unconditioned stimuli for the testing of adults and children.

One highly challenging observation of an informal and unsubstantiated nature concerned an apparent inverse

relationship between GSR and delay distortion. It was observed that for some individuals who could read under delayed feedback at a high level without apparent disturbance, marked skin resistance responses obtained. Although no systematic study of this has been carried out, a brief pilot experiment was done. For two highly sophisticated (trained) speaker-listeners both rate and GSR measures were obtained. It was noted that very little overt speech disturbance could be observed in either case but GSR measures were extremely indicative of a hearing response. Two other subjects were chosen who showed very severe overt speech disturbance in response to the delayed sidetone. The GSR recordings were not nearly as conclusive as the speech-rate distortion and not nearly as conclusive in these cases as for the other two speakers. The hypothesis is certainly suggested here that speakers are able to overcome the effects of delayed feedback only by a strength of effort which is measurable in terms of GSR. On the other hand, if the effort involved is admitted, support is brought to Abel's (*9*) contention that GSR is an index of an attitude of the organism directed toward overcoming difficulty. Further study is advised.

Refinement in Methodology

As might be expected, listeners tend to adapt fairly rapidly to the delayed sidetone signal. Although no controlled observations were made there appeared to be a significant adaptation which occurred over a series of stimulus presentations at a rate dependent upon the frequency and spacing of such presentations. Because of this it was felt that a wider spacing of sidetone signals would make the test more sensitive.

To investigate this hypothesis, eight subjects were tested at a 30 db sensation level under the same procedures outlined in the original experiment, with the following exceptions: (1) only seven stimuli were presented, all of which were the 'real' sidetone signal; and (2) the intervals between the stimulus presentations were at least seven chart lines apart (53 seconds).

The results of the judges' ratings on the presence of delayed sidetone indicated that all eight subjects had received the stimuli, whereas in the original technique only 70% of the subjects at the 40 db level were so detected. It should be noted that the records of these latter eight subjects were interspersed with records from the original experiment so as to prevent judgemental bias.

For those interested in the clinical application of this technique as a method of detecting non-organic hearing loss, these findings tend to suggest that it is possible to improve the sensitivity of the technique employed in the main body of this experiment by presenting the delayed sidetone signals at greater intervals of time.

Summary

In a study of the skin resistance changes appearing as unconditioned responses to delayed sidetone, GSR measures were made during the experimental employment of the delayed sidetone test of malingering on 50 normal hearing subjects. These subjects were divided into five groups in order to measure skin resistance changes under five sensation levels of delayed feedback: 80, 60, 40, 20, and 0 db with reference to spondee thresh-

old. The graphic records of skin resistance change resulting from 10 eight-second, randomly spaced administrations of delayed feedback during the reading of a long prose passage, were analyzed in two ways: (1) measures were made of the latencies of response to the sidetone; and (2) expert judgements of the graphic records were made. The following were among the results obtained: (1) The variation of the response latency to the sidetone appeared to be strongly related to sidetone level. High levels of sidetone produced highly consistent variations in skin resistance. (2) Analysis of the data seemed to indicate the possibility of measurable skin resistance changes in response to sidetone levels as low as 20 db. Strong responses were noted at 40 db. (3) Judges were able, with nothing more than the graphic records and latency measures to guide them, to make highly accurate judgements as to the level of sidetone employed with each subject. Of 30 persons hearing sidetone at or above 40 db only two (as an average for the judgements of two observers) were thought to have heard nothing. Of the 10 persons who received sidetone at the 0 db level, none was judged to have heard sidetone. Of the total of 100 judgements made by the two observers, only eight were in error by more than 20 db.

References

1. Butler, R. A., and Galloway, F. T., Factorial analysis of the delayed speech feedback phenomenon. *J. acoust. Soc. Amer.*, 29, 1957, 632-635.
2. Doerfler, L. G., and Stewart, K., Malingering and psychogenic deafness. *J. Speech Dis.*, 11, 1946, 181-186.
3. Hanley, C. N., and Tiffany, W. R., An investigation into the use of electro-mechanically delayed sidetone in auditory testing. *JSHD*, 19, 1954, 367-374.
4. Lee, B. S., Effects of delayed speech feedback. *J. acoust. Soc. Amer.*, 22, 1950, 824-826.
5. Spilka, B., Relationships between certain aspects of personality and some vocal effects of delayed speech feedback. *JSHD*, 19, 1954, 491-503.
6. Tiffany, W. R., and Hanley, C. N., Adaptation to delayed sidetone. *JSHD*, 21, 1956, 164-172.
7. Watson, L. A., and Tolan, T., *Hearing tests and hearing instruments*. Baltimore: Williams and Wilkins Co., 1949.
8. Woodworth, R. S., *Psychology*. 4th ed. New York: Henry Holt and Co., 1940.
9. Woodworth, R. S., *Experimental Psychology*. New York: Henry Holt and Co., 1938.

Behavioral Rigidity In Adult Stutterers

Ervin Kapos

Lloyd S. Standlee

In one sense stutterers are distinguished from non-stutterers by their tendency to persevere in fixed patterns of speech responses, and their inability to change from one sequence of responses to another as readily as the latter. Perseveration in fixed response patterns is the definition of behavioral rigidity, as measured by the index of behavioral stereotypy suggested by Miller and Frick (*4*), and the question immediately arises whether the general behavioral rigidity of stutterers is significantly higher than that of non-stutterers.

Working from the definition of perseveration as the tendency of a set of neurons, once excited, to persist in this state of excitation autonomously, and using a perseveration index based on initial and terminal performance on an alternating series of paper-and-pencil tasks of various types, Eisenson (*2*) exhibited results indicating differences in perseveration between stutterers and non-stutterers. No differences were found in initial or terminal speed or accuracy alone.

In a recent study, Kapos and Fattu (*3*) showed that the behavioral rigidity of severely speech-handicapped children was significantly higher than that of normal children, at least when the observed behavior consisted of the operation of a sufficiently complex multiple-choice maze. It was pointed out in this connection that the results of research on manipulatable factors affecting behavioral rigidity may therefore have considerable significance for speech therapy.

Ervin Kapos (B.A., Indiana University, 1954) is a Ph.D. Candidate at Indiana University. Lloyd S. Standlee (Ed.D., Teachers College, Columbia University, 1955) is Assistant Professor of Education at Indiana University.

Methodology

The purpose of the two experiments reported here was to compare the behavioral rigidity of groups of adult stutterers and non-stutterers. The measure of behavioral rigidity was the index of behavioral stereotypy proposed by Miller and Frick (*4*), applied to responses obtained from operation of a two-switch electromaze and a four-switch electromaze.

In both experiments, the subjects were informed that the problem was to find, in as few trials as possible, the switch which would illuminate the light (or lights) located on the same panel as the switches. Though the solutions to the problems were randomly selected, the subjects were not informed of this. They were required to continue responding until 102 responses had been obtained for each subject on the two-switch electro-

maze, and 201 responses on the four-switch electromaze. These yielded 100 sequences of three responses over two choices, and 200 sequences of two responses over four choices, respectively, —a sufficient number, as suggested in a recent paper by Kapos and Fattu (*3*), to establish asymptotic probabilities for calculation of third-order and second-order indices of behavioral stereotypy, respectively. More complete descriptions of the electromazes and of the index of behavioral stereotypy are also presented in the paper by Kapos and Fattu (*3*).

Subjects

The subjects in the experiment utilizing the two-switch electromaze were 30 students at Indiana University during the 1954-55 academic year. Fifteen were identified by the University Speech Clinic as stutterers, while the remaining 15 had no evident speech defects. The group of stutterers included 13 males and two females, ranging in age from 18 years and three months to 30 years and eight months at the time of testing, with a median age of 21 years and eight months.

The subjects in the experiment utilizing the four-switch electromaze were 32 students at Indiana University during the 1956-57 academic year. Sixteen were identified by the University Speech Clinic as stutterers, while the remaining 16 had no evident speech defects. The group of stutterers included 13 males and three females, ranging in age from 17 years and 11 months to 29 years and four months at the time of testing, with a median age of 19 years and 10 months. In both experiments the samples of non-stutterers were matched to the samples of stutterers in age and sex.

Results

Third-order indices of behavioral sterotypy (C_3) were computed for all subjects tested on the two-switch electromaze, and second-order indices (C_2) were computed for all subjects tested on the four-switch electromaze. Before male and female stutterers can be pooled in a single sample, it is necessary to verify that they are not sampled from different populations with respect to behavioral rigidity. (This result is available for non-stutterers of comparable ages from an unpublished study by the present authors.) The use of a parametric statistic is inappropriate because of the small number of female subjects and because of a lack of homogeneity in the variances for male subjects and female subjects. However, use of the Mann-Whitney statistic (*1*) on the indices of stereotypy for the sample tested on the two-switch maze yielded a U of 6, for n=13 and m=2 (critical values at the 10% level for the two-tailed test: 2, 24). For the sample tested on the four-switch maze, the obtained value of U was 29, for n=13 and m=3 (critical values at the 10% level for the two-tailed test: 6, 33). These results provide the desired verification.

TABLE 1. Comparison of third-order indices of behavioral stereotypy for groups of stutterers and non-stutterers tested on the two-switch electromaze.

Group	*N*	*M*	*S. D.*	*t*
Stutterers	15	.133	.071	.91
Non-stutterers	15	.112	.048	

TABLE 2. Comparison of second-order indices of behavioral stereotypy for groups of stutterers and non-stutterers on the four-switch electromaze.

Group	*N*	*M*	*S. D.*	*t*
Stutterers	16	.124	.095	.10
Non-stutterers	16	.121	.070	

The values of C_3 for the group of stutterers ranged from .025 to .239, with a mean of .133 and a median of .113. The indices for the group of non-stutterers ranged from .060 to .219, with a mean of .112 and a median of .100. The mean C_3 index for the group of stutterers was thus higher than that for the group of non-stutterers by .021, but a *t* test between these means (Table 1) shows the difference to be well short of significance at the five per cent level.

The values of C_2 for the group of stutterers tested on the four-switch electromaze ranged from .033 to .461, with a mean of .124 and a median of .110. The indices for the group of non-stutterers ranged from .028 to .259, with a mean of .121 and a median of .107. It is clear that these means do not differ significantly, and this is borne out by a *t* test between them (Table 2).

It must therefore be concluded that the foregoing studies offer no evidence that the behavioral rigidity of adult stutterers is higher than that of non-stutterers, at least as measured by the index of behavioral stereotypy computed for responses obtained in a two- or four-choice problem situation.

Summary

In two separate experiments, samples of adult stutterers and non-stutterers, matched with respect to age and sex, were tested on a two-switch electromaze and on a four-switch electromaze. No significant differences were observed in third-order indices of behavioral stereotypy obtained from responses to the two-switch maze, or in second-order indices of behavioral stereotypy obtained from responses to the four-switch maze. Thus the greater perseveration of stutterers in fixed speech patterns is seen not to imply a greater general behavioral rigidity than that of non-stutterers.

Acknowledgments

The authors are indebted to Dr. Nicholas A. Fattu, Director of the Indiana University Institute of Educational Research, for making apparatus and assistance available, and to Dr. Dean E. Williams of the Indiana University Speech Clinic for his cooperation in obtaining the subjects.

References

1. AUBLE, D., Extended tables for the Mann-Whitney statistic. *Bull. Inst. educ. Res.*, Indiana Univ., Vol. 1 (2), 1953.
2. EISENSON, J., A note on the perseverating tendency in stutterers. *J. genet. Psychol.*, 50, 1937, 195-198.
3. KAPOS, E., and FATTU, N. A., Behavioral rigidity in speech-handicapped children. *JSHD*, 22, 1957, 707-713.
4. MILLER, G. A., and FRICK, F. C., Statistical behavioristics and sequences of responses. *Psychol. Rev.*, 56, 1949, 311-324.